

COMMENT

On May 14, 2001 this sequence version replaced gi:2982535
 gi:2982536 gi:298454 gi:2982554 gi:2982562 gi:2984489 gi:2982572
 gi:2982574 gi:4493931.

For more information about this sequence or the Malaria Project,
 see http://www.sanger.ac.uk/Projects/P_falciparum.

FEATURES

```

source
  1..86827
    /organism="Plasmodium falciparum"
    /strain="3D7"
    /db_xref="taxon:5833"
    /chromosome="3"
    /clone="MAL3P5"
  join(322..603,826..1023,1197..1301,1458..2942)
    /gene="PFC0575w, MAL3P5.1"
  join(322..603,826..1023,1197..1301,1458..2942)
    /gene="PFC0575w, MAL3P5.1"
  /note="PFC0575w (MAL3P5.1). Hypothetical protein, len: 689
  aa, possible signal sequence, revised: added new exon 2"
  /codon_start=1
  /product="Hypothetical protein, PFC0575w"
  /protein_id="CAB38969.2"
  /db_xref="GI:711065"
  /db_xref="SPTREMBL:O97258"
  /translation="MYLKNVYIISCFILDLGFSFLLKMKYKHNHNNKSVTFPL
  RSPQYIKRRKRSIKNVFKKQKPLFEENLKKGSFGLGFRNQDQYIDVIS
  NINLRIKQVTHKKKNEFTENIKQLLHCVSFKIDFKIINNLSYIIKHFQMSNIT
  VHSILNIOISERVKKDAENYLAHLFLKDDENTLFSMMHIMFFKSKQVIEIRD
  IKSKKKKNSLIYNLPICLVITYCYMCLLIRYISHLICFFPFFCFCCYNLIER
  IYECVDLIRKRYERNLYCEKKIKFKMDALIKMEINMKDDDLXENVYHVELLRC
  FTKMLNHNKNIIRSNYDNIINNDISIDKMYMNPIDVINNISLDEKKEQENP
  DDENLKLDYEQFLPNDNIKIYEDQPLYNIDNSNNNNNNNTMKNRHKIK
  DYNDDDDYDEKEDLYIQNIDDIYIKNTIGMKSLEEFKNQFIEQADIEFQNFLS
  NVNLDQHRGNDSNTEKTHKNTKNTKNGYDTTELQNMENNFIKKNIDNNISLD
  ISNNEIKIKLKNOSDEINLSDLIYERLRTKVLWYIOKIEYLPKYQYDIINEQ
  YPIIKNETVLDLNYGKIVNSPDVDSLEFKTIDSIDPNEKKNQNMENQKNSKY
  IIMFNM"
  misc_feature
    604..609
      /gene="PFC0575w, MAL3P5.1"
      /note="predicted splice donor sequence for exon 1 of
      PFC0575w (revised)"
  misc_feature
    816..825
      /gene="PFC0575w, MAL3P5.1"
      /note="predicted splice acceptor sequence for exon 2
      (revised of PFC0575w)"
  misc_feature
    1024..1029
      /gene="PFC0575w, MAL3P5.1"
      /note="predicted splice donor sequence for exon 2 of
      PFC0575w (revised)"
  gene
    complement(3354..6644)
  CDS
    complement(3354..6644)
    /gene="PFC0580c, MAL3P5.2"
    /note="PFC0580c (MAL3P5.2), Hypothetical protein, len:
    1097 aa, possible signal sequence, predicted using
    hexExon"
  /codon_start=1
  /product="Hypothetical protein, PFC0580c"
  /protein_id="CAB38968.1"
  /db_xref="GI:4493932"
  /db_xref="SPTREMBL:O97257"
  /translation="MLGLKRNFFLLYVPSLFAFLKRRKDNENNYETLINNDIE
  KIKIRIHKCSIYPLLEFLNIYDSIYKILRLWYIFKRRKDKKEYYITNNVRK
  KRRAIKYTFNDEQNLKFTYIYVLEYSKYGILSPHLSLYILKIKSEHCVNYP
  SLIYNYKLDKHLNLEKKLAFKQIINHTQAPTNTHTNNNNNNKPKLDINHSC
  KNTNTSSYTNNMKENINIDYKYNHNFYKTSISKDNCQHTILNMILNLOTY
  DNICICLNTNININFYEMINILAYICYKMEIILLVYNHLEDKMKIKNOKNNTS
  LFKYIISFFKKEENHIYDLFDQMHLKKKDKNDKFNYSNETHNNIYKYSIDN
  FYDHNSSNRCSFKNLKQDDTKHIMGKEYIPMKSDEKKNNTNCGINIEK
  DQDKDLIKLYFGKNGKLDIQILNELXYMYMRLIFCSSLIKIKNHLLEKME
  FDKDKIILYLSADYNLNRRLRILKFSMEERENINSPASPFLLSKNIIFYEDEIG
  RSRDNTIYNNYDKETNTTNNNNNDNNTCSNNNDHICSNNNDHICSN
  NNNNICSNNNNICSNNNNCSNNKMLDEFCODNKFNDYNTKKEKRI
  YELAKIYTNNTIDYLGKKERHQEDNTINLIYIKKFPWIFLYLKNILKNTDSFIEH
  NNIVNGDINKNNLIFFKKYLVFESSIISFYIKDIYEVNKLRLYIYDNLKKFCR
  YFLKXNHEINKLIMKRAHRYIYNFDQFIINNYIHHKNNKHKHIHLKQCKDKE
  IDIYKDLTYCINNININISFYIHKVDHNECVYIRAFKNNKILLIETYNLNEKENIY
  YKKIKKYLAINNISNDLYPNISYKYNQYKNNKRNFSHFYSLKNDIHLLEF
  LYTORIONCDDISYIYKYNFNKPNPLNLYYELHYIVYSEKKKKKFFSISSSP
  XSDYMNSTFTSYFFSLSVLLFVHPDMYASYIFFKTLTYSGLPTYYYSLYNNI
  MVCGPKT"
  join(7669..7717,7800..7829,7912..7940,8064..8106,
  8302..8368,8568..8641,8713..8812,8924..9006,9122..9258,
  9369..9505,9613..9838)
    /gene="PFC0581w"
  join(7669..7717,7800..7829,7912..7940,8064..8106,
  8302..8368,8568..8641,8713..8812,8924..9006,9122..9258,
  9369..9505,9613..9838)
    /gene="PFC0581w"
  /note="PFC0581w, hypothetical protein, len: 324 aa,
  revised: new gene prediction, splicing prediction very
  tentative"
  /codon_start=1
  /product="Hypothetical protein, PFC0581w"
  /protein_id="CAB90285.1"
  /db_xref="GI:7711066"
  /translation="MKKKKHVFSIKFVNDKNKYLKSKDIIRINMIGLHIFRL
  QKJHIVANESQIASPPESFKNRYVTREDIDGIEKWFNITDYLSIESKIDNDFS
  NDIKTYTAMKNDQDLTTYSNNKIDTVNASQWQSPVFIKFSHWSPPGA
  LKVDDEIVSKNNLFSFALSNDNSVTKKIYDITLNDLIESETKYNFASVGVVY
  TLKKEKKIWRLLSKBKYPNMQVMDMKRKHITITVTINLFLSLSLSHRYHDSV
  QNFLEKKNNSDKLQDDIDEDEKYFDEILREAKKSEYDKDDDEEL"
  7790..7799
    /gene="PFC0581w"
    /note="potential splice acceptor sequence for exon 2 of
    PFC0581w"
  7830..7835
    /gene="PFC0581w"
    /note="potential splice donor sequence for exon 2 of
    PFC0581w"
  7901..7911
    /gene="PFC0581w"
    /note="potential splice acceptor sequence for exon 3 of
    PFC0581w"
  7941..7947
    /gene="PFC0581w"
    /note="potential splice donor sequence for exon 3 of
    PFC0581w"
  8055..8063
    /gene="PFC0581w"
    /note="potential splice acceptor sequence for exon 4 of
    PFC0581w"
  8107..8112
    /gene="PFC0581w"
    /note="potential splice donor sequence for exon 4 of
    PFC0581w"
  8293..8301
    /gene="PFC0581w"
    /note="potential splice acceptor sequence for exon 5 of
    PFC0581w"
  8369..8374
    /gene="PFC0581w"
    /note="potential splice donor sequence for exon 5 of
    PFC0581w"
  8559..8567
    /gene="PFC0581w"
    /note="potential splice acceptor sequence for exon 6 of
    PFC0581w"
  8642..8647
    /gene="PFC0581w"
    /note="potential splice donor sequence for exon 6 of
    PFC0581w"
  8704..8712
    /gene="PFC0581w"
    /note="potential splice acceptor sequence for exon 7 of
    PFC0581w"
  8813..8818
    /gene="PFC0581w"

```



```

/gene="PFC0581w"
/note="potential splice donor sequence for exon 7 of
PFC0581w"
misc_feature
8913..8923
/gene="PFC0581w"
/note="potential splice acceptor sequence for exon 8 of
PFC0581w"
misc_feature
9007..9012
/gene="PFC0581w"
/note="potential splice donor sequence for exon 8 of
PFC0581w"
misc_feature
9112..9121
/gene="PFC0581w"
/note="potential splice donor sequence for exon 9 of
PFC0581w"
misc_feature
9259..9264
/gene="PFC0581w"
/note="potential splice donor sequence for exon 9 of
PFC0581w"
misc_feature
9361..9368
/gene="PFC0581w"

Query Match          7.8%; Score 135; DB 3; Length 86827;
Best Local Similarity 46.3%; Pred. No. 1.3e-07;
Matches 585; Conservative 0; Mismatches 670; Indels 8; Gaps 4;

Qy 432 tttttttttcatcgtaacaaatctaaagatcggtatataaagaatcttcacaaaaa 491
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 39311 TTTATTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 39370

Qy 492 aaattggttagatggagtagccaaatttaacaaatcgtaaaacaaataaaacgatcgt 551
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 39371 ATAATTATTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 39430

Qy 552 agacaaatcctaagca-tcgtgcacaaagaatttaaaaaaatcgttttagtcaaatctaa 610
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 39431 TAATATATATATTTATTAAATAAATAAATAAATAAATAAATAAATAAATAA 39490

Qy 611 acaattgtataacaaatctaaacgtagaattgaaataaataaataaataaataaataa 670
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 39491 ATAATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 39550

Qy 671 tccaaatttaaatgacaaatctaaacgtagatgataacaaatctaaacgtagatgata 730
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 39551 TAAATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 39610

Qy 731 aatctaaatgatcgtacaaatattatgacattgttggcagggtggtgacggaa 790
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 39611 TAATAATTCAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 39670

Qy 791 cattttgtatattttctattatgggtttgtagaatttttctatttcgaaattgtctat 850
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 39671 TAAATGTTTTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 39730

Qy 851 acatatataataataattttaccattgttattatttttgcgaagaagcccttaataaa 910
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 39731 TATATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 39790

Qy 911 attgaattcgcataataaataatttttcccaaaaaaagaatagactgtctatctaaaa 970
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 39791 ATAATAATTTTATTTATTCACCTAATAATAAATAAATAAATAAATAAATAAATAA 39850

Qy 971 atttgattcccaatataagacaaatt----ctcaaaatgacaaacatttgaattctcg 1026
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 39851 TAATATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 39910

Qy 1027 atatagaacacatttctatttttgaattgggacataatccaaagttttatccaaacgta 1086
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 39911 TTTTAATTCATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 39970

Qy 1087 acttgaggaaaagtgttgattgagattacatccatattttttgttttttcattgaaattc 1146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 39971 AAATTAATGTTTACTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 40030

```

```

Qy 1147 atggaaaattaaatgcacacaaatgatgatgatgagattaaacaaagttttctcgttatt 1206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 40031 TATAATAATTATCATATATTTATTTATTTAAATTAATTTAAATTAATTAATTAATTAAT 40090

Qy 1207 gaattcttttttaaaaaaaccaacaaatctttaaactgtttgcaatagacaaatag 1266
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 40091 AATTATATAATATATATAGTATTTTAAATTTATTTATTTATTTATTTATTTATTTATTT 40150

Qy 1267 ttaatccatcgtggtctattgttagtaaaattgttaattttgtttatattaataaatt 1326
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 40151 TTTAATTATTCGAATACAAATATTAATAAATAAATAAATAAATAAATAAATAAATAA 40210

Qy 1327 ttgattttattgtatattgtatttagataacaaataaattgaatttaataatttttta 1386
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 40211 TTATTAAATCTGTTTTTAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAAT 40270

Qy 1387 tatcttaataaataacatttggtaattt-ttctatttttagaccatttctctatttttat 1445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 40271 TATTATAATATAATATATATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 40330

Qy 1446 ataacatttttaataactaaatgatgtgacacacacataattattttttatccaaagaaa 1505
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 40331 ATTATAATATAATATATATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 40390

Qy 1506 taatgctataaataatgggtctcttttatacacttcacataaattatgaataaaaaata 1565
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 40391 TTAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 40450

Qy 1566 aaatttaattataaattcttctcatcctcaatcgtaacagctagattcactatcaaca 1625
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 40451 AATTAAATATATTTAAATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 40508

Qy 1626 acttgggtatacaaaaggcgaagaattaaagcattatcgtgtgagccactttttctat 1685
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 40509 ATTTTATTTAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 40568

Qy 1686 cta 1688
    ||
Db 40569 TTA 40571

```

RESULT 6

```

AP001254/c LOCUS
DEFINITION Homo sapiens genomic DNA, chromosome 21q21.1-q21.2, clone:866B12,
LL56-APP region, complete sequence.
ACCESSION AP001254
VERSION AP001254.2 GI:7262571
KEYWORDS HTG.
SOURCE Homo sapiens DNA, clone:866B12.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 173915)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Tokoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Published Only in DataBase (2000) In press
JOURNAL Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Watanabe,H. and Sakaki,Y.
AUTHORS Direct Submission
TITLE Submitted (22-FEB-2000) to the DDBJ/EMBL/GenBank databases.
JOURNAL Masahira Hattori, The Institute of Physical and Chemical Research
(RIKEN), Genomic Sciences Center (GSC), Kitasato Univ., 1-15-1
Kitasato, Sagamihara, Kanagawa 228-8555, Japan
(E-mail:hattori@gs.c.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-42-778-9923, Fax:81-42-778-9924)
COMMENT On Mar 17, 2000 this sequence version replaced gi:7077189.
Sequence updated (15-Mar-2000).
FEATURES
Location/Qualifiers
source
1..173915

```


Hennig,S., Riesselmann,L., Dagand,E., Wehrmaeyer,S., Borzym,K.,
Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and
Yaspo,M.L.

TITLE JOURNAL

Direct Submission

Submitted (10-APR-2000) to the DDBJ/EMBL/GenBank databases. The
Chromosome 21 Mapping and Sequencing Consortium: * RIKEN Genomic
Sciences Center, Human Genome Research Group * Institute of
Molecular Biotechnology, Genome Analysis * Keio University School
of Medicine, Dept. of Molecular Biology * GBF, Dept. of Genome
Analysis * Max-Planck Institute for Molecular Genetics (addresses
see below)

On May 30, 2000 this sequence version replaced gi:7717275.

The chromosome 21 mapping and sequencing consortium consisting of
* RIKEN Genomic Sciences Center, Human Genome Research Group, *

Sagamihara 228-8555, Japan,

* e.mail: hattori@gsc.riken.go.jp

* URL: http://hgp.gsc.riken.go.jp/

and

* Institute of Molecular Biotechnology, Genome Analysis, *

Beutenbergstrasse 11, D-07745 Jena, Germany,

* e.mail: gscj-submit@genome.imb-jena.de

* URL: http://genome.imb-jena.de/

and

* Keio University School of Medicine, Molecular Biology, * Tokyo
160-8582, Japan,

* e.mail: nshimizuedmb-med.keio.ac.jp

* URL: http://www.dmb.med.keio.ac.jp/

and

* GBF, Dept. of Genome Analysis,

* Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e.mail:
info.genome@gbf.de

* URL: http://genome.gbf.de/

and

* Max-Planck Institute for Molecular Genetics,

* Ihnestrasse 73, D-14195 Berlin, Germany,

* e.mail: info-chr21@molgen.mpg.de

* URL: http://chr21.rz-berlin.mpg.de/

AL163223: Submitted (10-APR-2000).

Location/Qualifiers

1. 340000

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="21"

<1. 14692

/note="Accession No. AF238375"

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q21.1"

/clone="R49B5, 5' partial"

/clone_lib="RPCI-11 BAC library"

<1. 121279

/note="Accession No. AP001256"

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q21.1"

/clone="R732H1, 5' partial"

/clone_lib="RPCI-11 BAC library"

1417. 1451

/note="(TG)n"

/rpt_family="Simple_repeat"

/rpt_type=TANDEM

1731. 1890

/note="LTR33"

/rpt_family="LTR/Retroviral"

/rpt_type=DISPERSED

1891. 2300

/note="MLT1B"

/rpt_family="LTR/MaLR"

/rpt_type=DISPERSED

2301. 2431

/note="LTR33"

repeat_region
/rpt_family="LTR/Retroviral"
/rpt_type=DISPERSED
complement(2585.. 2863)
/note="AluSq"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
2946.. 3049
/note="(TTATA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
3202.. 3581
/note="THEIC"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
4341.. 4632
/note="AluSq"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(7166.. 7371)
/note="MIR"
/rpt_family="SINE/MIR"
/rpt_type=DISPERSED
7454.. 7717
/note="AluSc"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(7845.. 8207)
/note="L1ME1"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
8302.. 8478
/note="(TA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
complement(8485.. 9080)
/note="L1ME1"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
complement(9143.. 9442)
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
9559.. 9766
/note="MLT2E"
/rpt_family="LTR/Retroviral"
/rpt_type=DISPERSED
complement(9852.. 9961)
/note="L1ME1"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
complement(9962.. 10247)
/note="AluY"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(10248.. 10367)
/note="L1ME1"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
10713.. 10821
/note="L1MEC"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
complement(11480.. 11662)
/note="MER5A"
/rpt_family="DNA/MER1_type"
/rpt_type=DISPERSED
11744.. 11946
/note="MER8"
/rpt_family="DNA/MER2_type"
/rpt_type=DISPERSED
complement(13505.. 13806)
/note="AluX"
/rpt_family="SINE/Alu"

FEATURES source

source

source

repeat_region

repeat_region

repeat_region

repeat_region

RESULT	8
AC066608	
LOCUS	AC066608
DEFINITION	172805 bp DNA Homo sapiens chromosome 3 clone RP11-785A7 map 3p, complete sequence.
ACCESSION	AC066608
VERSION	AC066608.5
KEYWORDS	GI:12745084
SOURCE	HTG. human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 172805) Xiong,H., Zhou,Y., Dong,H., Lin,W., Chen,B., Zhang,C., Zhang,Y., Cai,L., Ying,H.F., Wang,H., Gu,W., Zhu,G., Tu,Y., Zhang,X., Jia,J, Shen,H., Zhang,D., Wu,C., Lu,G., Zhong,M., Jiang,H. Ren,S., Fu,G

QY	1544	gataattatgaataaaataaaaatttaattatatataaatcatttcacgtcaaa 1503 Db 141431 TATATTATATAAAATATATATATATATTTTATATATATATATATATACCA 141490					
QY	1604	gctagattactactaacacaaccttggtagataaaaaaggcgagaagt 1652 -					
Db	141491	TATTTATATATTCATCTATATATATATATATAGATAAATATAT 141539					
RESULT	9						
AC046181		DNA	HTG	30-JUL-2001			
LOCUS	Homo sapiens chromosome 18 clone RP11-815K1 map 18, *** SEQUENCING IN PROGRESS ***, 2 ordered pieces.						
ACCESSION	AC046181						
VERSION	AC046181.6 GI:15028577						
KEYWORDS	HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.						
SOURCE	human.						
ORGANISM	Homo sapiens						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 172666)						
AUTHORS	Birren,B., Linton,L., Nusbaum,C. and Lander,E.						
TITLE	Homo sapiens chromosome 18, clone RP11-815K11						
JOURNAL	Unpublished						
REFERENCE	2 (bases 1 to 172666) Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bedalov,F., Boguslavsky,L., Bouckgalter,B., Brown,A., Burkett,G., Campiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Guyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., LaRocque,K., Lamarezes,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., MCPheeters,R., Meldrum,J., Menues,L., Mihova,T., Miranda,C., Mieng,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange,Thomann,N., Stanjovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodores,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.						
TITLE	Direct Submission						
JOURNAL	Submitted (13-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA						
COMMENT	On Jul 29, 2001 this sequence version replaced gi:14718366. All repeats were identified using RepeatMasker: Smith,A.F.A. & Green,P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html ----- Genome Center ----- Center: Whitehead Institute/ MIT Center for Genome Research Center code: WIBR Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@genome.wi.mit.edu ----- Project Information ----- Center project name: L8732 Center clone name: 815_K_11 -----						
*	NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submittor.						
*	This sequence will be replaced						
*	by the finished sequence as soon as it is available and the accession number will be preserved.						
*	The accession number will be preserved.						
*	1						

Qy	366	tctttctcttttctcttctttttttccgcgtgcgattcttcccatctgtctatcgtttttt	425
Db	68056	TCATATACATTTTATATAAATTTTCATATACATTTTATATAATATATATATAATTTT	68115
Qy	426	-tctcttttttttttcaactcgtaaccaaactcaaaagatcgatataaagaatcttca	484
Db	68116	ATACATATAAAATTTGTATATATAAATTTTAATATATAAAATTTATATAATTTTATAT	68175
Qy	485	aaaaaataattggttagattggagtagcgaatttaaacaaatcgctcaaaaaataaaa	544
Db	68176	ATAAAATATATAAATTTTAATATAAAATTTTATATAAAATTTTAAATATATATAAAT	68235
Qy	545	cgatcgtagacaaactcaaacgatcgctgcacaaagaatttaaaaaaatcgttagtcaa	604
Db	68236	ATATGAATTTTAAATATATAAAATTTATATATAAATTTTATATAAAATTTATGTATAA	68295
Qy	605	atctaaacaattgataaccaaact-taaacgatagaatcigaatacaataaactcggttagat	663
Db	68296	ATTATATGTATATTACATATAAAATATATATAAATATAAATATATATAAATATATATT	68355
Qy	664	ttggctatccaaattttaaatgaccaaactcaaacgatcgtaaccacaaactcaaacgatcg	723
Db	68356	TAATATAAATATACATAAATATATAATATAAATATACATTTATATTAAATATATATAA	68415
Qy	724	kataccaaactcaaatgatcatgtaccacaaatatattatgcacattgttggcaggggtggtt	783
Db	68416	TATATAAATAGATATATAAATATATATATATATATATATATATATATATATATATATA	68475
Qy	784	gaacgaaactttgtatatcttctattatgggttggtagaaatttttccatttccgaaatt	843
Db	68476	TATAAATATATTAAATAGATATATAAATATAAATATATCTATATAAATATCTATTATATA	68535
Qy	844	gtctatacaataataataataatttttaccacttcgattatttttcgaaaaagaccctt	903
Db	68536	TAAATATATAAATATAATATACATATAAATATAGATACATTTATATAAATATAATATCA	68595
Qy	904	taataaattgaaatcgcataaatttaaaatttttcccaaaaaagtagacttgtcta	963
Db	68596	TATATAATATCTATT--ATATAATAAATAATATCTATATATAAATCTATTATATATA	68652
Qy	964	tctaaaaatttggattcccaatatagaacaaattctcaaaaatgcacaaacatttgaattc	1023
Db	68653	ATAAATATATCTATATAAATATCTATTATATAAATAAATATATCTATATAAATATAT	68712
Qy	1024	tcgatatagaaaaacttactattttgaaattgggacatttccaaagtcttcttccaaac	1083
Db	68713	CTATTATATAAATAGATATATATATCTATT--ATATAATAAATAAATATATCT	68766
Qy	1084	gtaactttgaaggaaaagttagtgagattcacatccattttgttttccatttgaatt	1143
Db	68767	ATTATATATAAATATAAATATATATATAAATTTTATATTTTTAATATATAAATAAAT	68826
Qy	1144	tctatgaaaaataaaatgcacacaaaaatgatgtatgagattaaaccaaagtcttactcgtt	1203
Db	68827	ATAAATAAATAATTTTATAAATAAATAAATATATATTTTATATATTTTATATATATAT	68886

	repeat_region	/rpt_family="MIR3"	complement(46397. .46526)	
	repeat_region	/rpt_family="L2"	complement(46565. .46820)	
	repeat_region	/rpt_family="MIR"	complement(47841. .48058)	
	repeat_region	/rpt_family="L2"	.48505. .48525	
	repeat_region	/rpt_family="(TPCA)n"	complement(49027. .49257)	
	repeat_region	/rpt_family="L2"	complement(49283. .49426)	
	repeat_region	/rpt_family="MIR"		

Query Match 7.7% Score 133.6; DB 9; Length 176174;
Best Local Similarity 46.4%; Pred. NO. 1.7e-07;
Matches 620; Conservative 1; Mismatches 695; Indels 20; Gaps

Qy	366	tcttttcttttctcgttttttttttttcgcgtggagattctcccaatgcttatcggttttt	425
Dy	68056	TCTATATACATTTTATAATATTTTCATATACATTTTTATAAATATATATATAATTTTT	68115
Qy	426	-tctcttttttttttcacatcgtaaccaaatctcaaagatcgtataaaagaactctca	484
Dy	68116	ATACATATAAAATTTGTATATATAATTTTATATATAAATTTATATAATTTTAATAT	68175
Qy	485	aaaaaaaattgttttagattggagtgcgaatttaaacaatcgcgtaaaaaaataaa	544
Dy	68176	ATAAAATATAAAATTTTATATAAAAATTTATATAAAAAATTTTAAATATATAAATTTAT	68235
Qy	545	cgatcgtagacaaatctaaacgatcgtagcaaaaagatttaaaaaaatcgtttagtcaa	604
Dy	68236	ATATGAATTTTAAATATATAAATTTATATAAATTTTATATAAATAAATTTATGATATAA	68295
Qy	605	atctaacaatttgataaaccaat-taaacgatagaattgaaatacaaaatcggttagat	663
Dy	68296	ATTATATGTTATTTACATATAAATATATAAATATAAATATAAATATAAATATAATTTATA	68355
Qy	664	tggctaccaaatttaagtaccaaactctaaacgatcgtagcaaaatctcaaacgatcg	723
Dy	68356	TAATATAAATACATAAATATATAAATATAAATACATTTATATTAATATAATATAA	68415
Qy	724	kataccaatatgaatgatcatgtaccaaaatatattgcacattgttggcagggttggct	783
Dy	68416	TATATAAATAGATATAAATATATATATATATATATATATATATAAATATATATATA	68475
Qy	784	gaocgaacatttgttatattctattatgggttggtagaatttttcatttttcgaaatt	843
Dy	68476	TATAAATATATTAATAGATATATAAATAAATATATCTATATAAATATCTATATAATA	68535
Qy	844	gtctatacaatakaatatataattttcacacttgattatttttcgaaaagaccct	903
Dy	68536	TAAATATATAAATAATATATCAATATAAATAGATACATTTATATAAATAAATAATCAA	68595
Qy	904	taataaattgaattcgcataataataaaattttttccccaaaaagtagactatgtcta	963
Dy	68596	TATATAATATCTATT--ATATAAATAAATAATATCTATATAAATAATATCTATATATA	68652
Qy	964	tctaaaaatttgattcccacatatagaacaattctcaaaatgacaaacatttgaaattc	1023
Dy	68653	ATAAATATCTATATAATATCTTTATATAAATAAATAATATCTATATAATATAATAT	68712
Qy	1024	tcgatagaacaaccttactattttgaaattggacatatatccaaagtttattccaaac	1083
Dy	68713	CTATTATATAAATAGATATATATATCTATT-----ATATAAATAAATAAATAATCT	68766
Qy	1084	gtaactttgaaggaaaagttgattggagattcacatccatatttttggtttttcattgaa	1143
Dy	68767	ATTATATAAATAATATAATATGTAATATAAATTTTATATTTTAAATATAAATAAATAA	68826
Qy	1144	ttaatgaaaaattaaaatgcacacaaaatgatgtatgagattaaaccaaagtttcatg	1203
Dy	68827	ATAAATAAATAATTTTATAAATAAATAAATAATTTTATAATTTTATAATTTATATATAT	68886

QY 1204 attgaattctttttattataaaaa-----ccaaacaaaatttttaaaactgttgcgaat 1254
Db 68887 GATATATTTATATATATATATATATATTTTCTATACAAATATATATACAAATTTTATATAT 68946
QY 1255 agaccaataatagtttaaccatcggtctctattgttagaataaactgtaattttgttatat 1314
Db 68947 AAATATATATATATATATATATATATATATTTTATATATATTTCTATATTTTATATATAA 69006
QY 1315 ttaataaaatattttgattttattttgatatattttgatttttagataacaaaataagattta 1374
Db 69007 ACTTATATTTAT 69066
QY 1375 aatattattttatattcttaataataaacattgttgaattttttctatttttagaccatttct 1434
Db 69067 AAATTTATTTCTCTACA 69126
QY 1435 cttatttttataacattttaactaaatgaatgatgcacacacactaattattttttta 1494
Db 69127 TAAATTTTA 69186
QY 1495 tccaaagaaaatagctataaaatattgggtcttctttatcccttcattgataaattatga 1554
Db 69187 TATTTTA 69246
QY 1555 aaaaataaaataaaatttaattataataattcatttcattcattcgcacagctagattatta 1614
Db 69247 TATTA 69306
QY 1615 ctatatcaaacacttggatataaaaggcgaagaaataagcattatcggtgagccac 1674
Db 69307 TTTAA 69366
QY 1675 tttttcttatctaga 1690
Db 69367 TGTATATATTTATATA 69382

RESULT 11
AC093082/c
LOCUS AC093082 326924 bp DNA HTG 09-AUG-2001
DEFINITION Homo sapiens chromosome UNK clone RP11-505B3, *** SEQUENCING IN
PROGRESS ***, 42 unordered pieces.
ACCESSION AC093082
VERSION AC093082.1 GI:15144475
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 326924)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 326924)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (09-AUG-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
----- Project Information -----
Center project name: H.NH0505B03

* NOTE: This is a 'working draft' sequence. It currently
* consists of 42 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1568: contig of 1568 bp in length
1668: gap of unknown length
3345: contig of 1677 bp in length
3346: gap of unknown length
3445: contig of 1480 bp in length
4925: gap of unknown length
5026: contig of 2227 bp in length
7253: gap of unknown length
7352: contig of 1838 bp in length
9190: gap of unknown length
9290: gap of unknown length
9291: 1018: contig of 1728 bp in length
11019: 1118: gap of unknown length
11119: 13028: contig of 1910 bp in length
13029: 13128: gap of unknown length
13129: 14658: contig of 1530 bp in length
14659: 14758: gap of unknown length
14759: 16766: contig of 2008 bp in length
16767: 16866: gap of unknown length
16867: 19394: contig of 2528 bp in length
19395: 19494: gap of unknown length
19495: 22382: contig of 2888 bp in length
22383: 22482: gap of unknown length
22483: 26670: contig of 4188 bp in length
26671: 26770: gap of unknown length
26771: 28899: contig of 3129 bp in length
28899: 29999: gap of unknown length
29999: 32393: contig of 2394 bp in length
32394: 32493: gap of unknown length
32493: 35161: contig of 2668 bp in length
35162: 35261: gap of unknown length
35262: 39253: contig of 3992 bp in length
39254: 39353: gap of unknown length
39354: 42753: contig of 3400 bp in length
42754: 42853: gap of unknown length
42854: 45570: contig of 2717 bp in length
45571: 45670: gap of unknown length
45671: 50008: contig of 4338 bp in length
50009: 50108: gap of unknown length
50109: 55872: contig of 5764 bp in length
55873: 55972: gap of unknown length
55973: 60627: contig of 4655 bp in length
60628: 60727: gap of unknown length
60728: 64794: contig of 4067 bp in length
64795: 64894: gap of unknown length
64895: 69682: contig of 4788 bp in length
69683: 69782: gap of unknown length
69783: 74782: contig of 5000 bp in length
74783: 74882: gap of unknown length
74883: 80238: contig of 5356 bp in length
80239: 80338: gap of unknown length
80339: 86095: contig of 5757 bp in length
86096: 86196: gap of unknown length
86196: 93489: contig of 7294 bp in length
93490: 93589: gap of unknown length
93590: 99512: contig of 5923 bp in length
99513: 99612: gap of unknown length
99613: 105968: contig of 6356 bp in length
105969: 106068: gap of unknown length
112079: contig of 6011 bp in length
112179: gap of unknown length
118180: 118180: contig of 6001 bp in length
118181: 118280: gap of unknown length
118281: 126420: contig of 8140 bp in length
126421: 126520: gap of unknown length
134015: contig of 7495 bp in length
134115: gap of unknown length
134116: 143112: contig of 9297 bp in length
143113: 143512: gap of unknown length
143513: 151892: contig of 8380 bp in length
151893: 151992: gap of unknown length

FEATURES

```
1. 326924
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="UNK"
/clones="RP11-505B3"
1. 1568
/note="assembly_name:Contig10"
1669. 3345
/note="assembly_name:Contig11"
3446. 4925
/note="assembly_name:Contig13"
5026. 7252
/note="assembly_name:Contig15"
7353. 9190
/note="assembly_name:Contig16"
9291. 11018
/note="assembly_name:Contig17"
11119. 13028
/note="assembly_name:Contig18"
13129. 14658
/note="assembly_name:Contig20"
14759. 16766
/note="assembly_name:Contig21"
16867. 19394
/note="assembly_name:Contig22"
19495. 22382
/note="assembly_name:Contig23"
22483. 26670
/note="assembly_name:Contig24"
26771. 29899
/note="assembly_name:Contig25"
30000. 32393
/note="assembly_name:Contig26"
32494. 35161
/note="assembly_name:Contig27"
35262. 39253
/note="assembly_name:Contig28"
39354. 42753
/note="assembly_name:Contig29"
42854. 45570
/note="assembly_name:Contig30"
45671. 50008
/note="assembly_name:Contig31"
50109. 55872
/note="assembly_name:Contig32"
55973. 60627
/note="assembly_name:Contig33"
60728. 64794
/note="assembly_name:Contig34"
64895. 69682
/note="assembly_name:Contig35"
69783. 74782
/note="assembly_name:Contig36"
74883. 80238
/note="assembly_name:Contig37"
80339. 86095
/note="assembly_name:Contig38"
86196. 93489
/note="assembly_name:Contig39"
```


QY	1127	tgttttcatattgaatttcattggaataattaaaatgcacacaaaatgatgatgagatta	1186
Db	73785	TAATTGTGTAATAATTAATTAATAAACAATTTATTAAATTAATTAATATATAAATAATTA	73722
QY	1187	aaccaaggtttatcggtattgaattcttttatcaaaaaaacacaaaattttaaaacttg	1246
Db	73725	TATATTATCTTAATAATTAATTAATTTTAAATTTATTTTAATAATAAATTAATTTATTAAT	73666
QY	1247	tttgcaatagacccaatatagtttaatccatcgctggctcattgttagataaaatttgaataattt	1306
Db	73665	ATTAATAATGCTATATTTTAAATTTGTTTAAATTTATTTATTTATTTATTTATTTATTTAT	73607
QY	1307	tggtatttttaataaataatttttgattfattttgatataattgttatttagataacaaaatt	1366
Db	73606	TATTAATTTAAATTAATTTATTTATTTATTTAAATTTATTTATTTATTTATTTATTTATTT	73547
QY	1367	agagtttaaatattatttttatatactcttaataataaacacattttgttaatttttctattttaga	1426
Db	73546	AATTT-----AATTAATTTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA	73492
QY	1427	ccatttctcttttttataacatttttaactaaatgaatgatgacacacactaatat	1486
Db	73491	TTTATTATATATATATATTTAATTAATTTATATAATAATTAATTAATTAATTAATTTAATTT	73432
QY	1487	tattttatccaaagaaaataatgctataaaaatggctgtctctttatcacaccttcacgat	1546
Db	73431	TATTTATATTAATTAATTAATTAATTAATTTATTTAAATTTAATTAATTAATTTATTTATTA	73372
QY	1547	aattatgaaaaataaaaaaatttaattataatcatttcatttcctaatcgtgtacaagct	1606
Db	73371	TTTATTTTATTTAATTAATTTAAATTTAATTTATATATAT-ATTTAATTTAAATTAATTAATTA	73313
QY	1607	agatatcactatatacaaacatttggctatataaaaggcggaagaataaagcattatcgtg	1666
Db	73312	TTTAATTTAAATTTATATATATATTAATTTAAATTTAATTTATATATATTTAATTTATTTATATAT	73253
QY	1667	tgagccacttttctatatctagagatagaagggtttaaaacatgctctcctaatctggaaa	1725
Db	73252	TTATTTAATTTATTTATATTTAATTTAATTTATATATATTTATTTATTTATTTATTTATTTA	73194
RESULT	13		
AC004157/c			
LOCUS	AC004157	169546 bp	DNA
DEFINITION	Plasmodium falciiparum chromosome 12 clone 3D7, *** SEQUENCING IN PROGRESS ***, 2 unordered pieces.		
ACCESSION	AC004157		
VERSION	AC004157.8	GI:9797712	
KEYWORDS	HTG; HTGS-PHASEL.		
SOURCE	malaria parasite P. falciiparum.		
ORGANISM	Plasmodium falciiparum		
REFERENCE	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
AUTHORS	1 (bases 1 to 169546)		
TITLE	Hyman, R.W., Fung, E.L., Qin, F., Rowley, D., Mao, J., Tamaki, T.,		
REFERENCE	Kurd, O.B., Conway, A.B. and Davis, R.W.		
AUTHORS	Plasmodium falciiparum 3D7 chromosome 12		
TITLE	Unpublished		
REFERENCE	2 (bases 1 to 169546)		
AUTHORS	Hyman, R.W., Qin, F., Fung, E.L., Conway, A.B. and Davis, R.W.		
TITLE	Direct Submission		
JOURNAL	Submitted (19-FEB-1998) Stanford DNA Sequencing and Technology		
JOURNAL	Center, Stanford University, 855 California Avenue, Palo Alto,		

COMMENT

54304, USA
On 04/12/2000 this sequence version replaced g1:8810437.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

[illegible]

Qy	1015	ttgaattctcgtatagaaacaatttacttatttgtaattgggacatatattccaagt	1074
Db	85489	TTAATTAATAAAATATATATATATTTTTTAAATAATAATAATAATATTTCACATTTT	85430
Qy	1075	attccaaacgttaa-----ctttgaaggaaaagttgatggagattcacataatttt	1126
Db	85429	AATTAATAATAAAATAACCATTATTAATTAACCTTAATAATATAATAAATAATAATT	85370
Qy	1127	tgttttcattatgaattcatcgaaaaattaaaaatgcacacaaaatgatgtatgagatta	1186
Db	85369	TAAUUGTGTAATTAATAATAATAAAAAACATTTATTAATAATAATAATAATAATAATA	85310
Qy	1187	aaccaagaatttatcgatttgaattctttttttaaaaaaaccaaaaaatttttaaaacttg	1246
Db	85309	TATATATGCTTAATTAATAATAATTTTTTTTAAATATTTTAAATAATAATAATAATAAT	85250
Qy	1247	tttgcacatagaccacatatagttaaccatcggtctattgtagataaaatgtgaatttt	1306
Db	85249	ATTAATTATGTATATTTTATTTAATGGTTTAATATTATTTATTTATTT-TAATATAT	85191
Qy	1307	tgttatatttaataaaatatttttgtatttttgatataattttgtattttagatacaaaaa	1366
Db	85190	TATTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	85131
Qy	1367	aagatttaaaatcatattttatctctaatacaaacatttgttaattttttctattttaga	1426
Db	85130	AATTT- ----AATTAATGCTTTTATTTATTAATTAATTAATTAATTAATTAATTAAT	85076
Qy	1427	ccattctctctattttatataaacatttttaatacaaatgaatgatgcacacacacaaat	1486
Db	85075	TTTATTAATATATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	85016
Qy	1487	tattttatccaaagaaaaaatgcatacaaaatgggtctctctttatcacaccttcgat	1546
Db	85015	TATTTATATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	84956
Qy	1547	aattatgaaaaataaaaaataatttaattataattcatctcatcgaatcgtaacagct	1606
Db	84955	TTTATTTATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	84897
Qy	1607	agattactatatacaaacactttgtgtcataaaaaagggcaagaaattgaacattatcgtg	1666
Db	84896	TTTAATTTAATATATATATATATTTAAATTAATTTAGTATATATATTTAATTTATATAT	84837
Qy	1667	tgagcccattttctatatctagagatagaagggtttaaaaatcagtcctctaatggaaa	1725
Db	84836	TTATTTAATTTATATATATATTTAATTAATTAATTAATTTATTTATTTTATTAACATAAA	84778
RESULT	14		
AC091742/c			
LOCUS	AC091742	188930 bp	DNA HTG
DEFINITION	Homo sapiens chromosome UNK clone RP11-1220K2, WORKING DRAFT SEQUENCE, 4 unordered pieces.		09-AUG-2001
ACCESSION	AC091742		
VERSION	AC091742.4	GI:15144365	
KEYWORDS	HTG; PHASEI; HTGS_DRAFT; HTGS_ACTIVEFIN.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 188930)		
JOURNAL	The sequence of Homo sapiens clone		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 188930)		
JOURNAL	Waterston,R.H.		
REFERENCE	Submitted Submission		
AUTHORS	Direct Submission		
JOURNAL	University School of Medicine, 444 Forest Park Parkway, St. Louis, MO 63108, USA		
COMMENT	On Aug 9, 2001 this sequence version replaced gi:14329170.		

Search completed: April 2, 2002, 08:49:04
Job time: 11391 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 2, 2002, 06:29:29 ; Search time 181.18 Seconds
(without alignments)
8209.843 Million cell updates/sec

Title: US-09-811-093-42
Perfect score: 1735
Sequence: 1 agcggatacaattccac.....taattggaagcttgtagt 1735

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues
Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	N_Geneseq_1101.*
1:	/SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	279	16.1	936	AAF58252	Oligonucleotide D1
C 2	279	16.1	936	AAF58254	Oligonucleotide D1
C 3	279	16.1	936	AAF58257	Oligonucleotide D1
C 4	279	16.1	936	AAF58259	Oligonucleotide D1
C 5	279	16.1	936	AAF58262	Oligonucleotide D2
C 6	279	16.1	936	AAF58255	Oligonucleotide D1
7	278.2	16.0	936	AAF58252	Oligonucleotide D1
8	278.2	16.0	936	AAF58254	Oligonucleotide D1
9	278.2	16.0	936	AAF58257	Oligonucleotide D1
10	278.2	16.0	936	AAF58259	Oligonucleotide D2
11	278.2	16.0	936	AAF58262	Oligonucleotide D2

C 12	278.2	16.0	938	22	AAF58255	Oligonucleotide D1
C 13	130	7.5	217	21	AAAG7340	Human UNK22-5' cDN
C 14	109.4	6.3	700	22	AAH93026	Human inflammatory
C 15	109	6.3	3244	20	AAH59168	Rat G protein coup
C 16	108.8	6.3	1272	19	AAV19124	Nucleotide sequenc
C 17	108	6.2	371	21	AAA70054	Human ovarian carc
C 18	107.6	6.2	1410	22	AAF91975	B thuringiensis 44
C 19	106.6	6.1	848	21	AAZ97334	Human prostate can
C 20	106.2	6.1	3908	22	AAH13152	Escherichia coli T
C 21	105.2	6.1	724	21	AAZ97430	Human prostate can
C 22	105.2	6.1	740	21	AAZ97429	Human prostate can
C 23	105.2	6.1	817	21	AAZ97370	Human prostate can
C 24	105.2	6.1	1024	21	AAZ97209	Human prostate can
C 25	105.2	6.1	1024	21	AAZ97219	Human prostate can
C 26	104.8	6.0	771	21	AAZ97308	Human prostate can
C 27	104.8	6.0	809	21	AAZ97398	Human prostate can
C 28	104.8	6.0	815	21	AAZ97377	Human prostate can
C 29	104.8	6.0	822	21	AAZ97362	Human prostate can
C 30	104.8	6.0	823	21	AAZ97319	Human prostate can
C 31	104.8	6.0	827	21	AAZ97356	Human prostate can
C 32	104.8	6.0	1017	21	AAZ97187	Human prostate can
C 33	104.8	6.0	1024	21	AAZ97159	Human prostate can
C 34	104.8	6.0	1024	21	AAZ97217	Human prostate can
C 35	104.8	6.0	2016	22	AAH03518	Mouse cell cycle c
C 36	104.8	6.0	2672	21	AAH30885	Rat GFAT II coding
C 37	104.4	6.0	198	22	AAH83749	Human ovarian tumo
C 38	103.2	5.9	1024	21	AAZ97239	Human prostate can
C 39	102.8	5.9	775	21	AAZ97272	Human prostate can
C 40	102.6	5.9	1021	21	AAZ97242	Human prostate can
C 41	102	5.9	435	22	AAH17792	Human breast cance
C 42	99.2	5.7	704	20	AAH90804	Rat progression re
C 43	99.2	5.7	837	21	AAZ97338	Human prostate can
C 44	99.2	5.7	5990	22	AAF80223	Nucleotide sequenc
C 45	99	5.7	779	21	AAZ97278	Human prostate can

ALIGNMENTS

RESULT 1	
AAF58252/c	
ID	AAF58252 standard; DNA; 936 BP.
XX	
AC	AAF58252;
XX	
DT	24-APR-2001 (first entry)
XX	
DE	Oligonucleotide D1835.
XX	
KW	Electron-transfer group; ETM; mismatch; genotyping;
KW	gene expression; ss.
XX	
OS	Synthetic.
XX	
PN	WO200107665-A2.
XX	
PD	01-FEB-2001.
XX	
PF	26-JUL-2000; 2000WO-US20476.
XX	
PR	26-JUL-1999; 99US-0145695.
PR	17-MAR-2000; 2000US-0190259.
XX	
XX	(CLIN-) CLINICAL MICRO SENSORS INC.
XX	
XX	Umek RM;
XX	
XX	WPI; 2001-159728/16.

Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface

[illegible]

RESULT 4
AAF58259/C
ID AAF58259 standard; DNA; 936 BP.
XX AC AAF58259;

[illegible]

D	b	542	www.....tag.....	601
Q	y	1423	tagaccattctcttatttttatataaacattttaataacaaatgaatgagcacacacta	1482
D	b	602	www.....	661
Q	y	1483	atatattttttccaaagaataatgcataaaaatgggtctctctttatcaccttca	1542
D	b	662	www.....	721
Q	y	1543	tgataattatgaaaaataaaaaatttaattataataattcatttcattcatctaatcgta	1602
D	b	722	www.....	781
Q	y	1603	a 1603	
D	b	782	w 782	
R	E	S	U	L
A	F	I	D	
I	D	I	D	
A	C	A	A	
X	X	D	T	
X	X	D	E	
X	X	K	W	
X	X	K	W	
X	X	O	S	
P	N	P	N	
P	D	P	F	
X	X	P	F	
X	X	P	R	
X	X	P	R	
X	X	(C	L	
X	X	U	m	
X	X	W	P	
X	X	N	u	
X	X	P	T	
X	X	P	T	
X	X	P	S	
X	X	C	C	
X	X	C	C	
X	X	C	C	
X	X	C	C	
X	X	C	C	
X	X	S	Q	

AAAF58257
ID AAF58257 standard; DNA; 936 BP.

AAF58257;
24-APR-2001 (first entry)
Oligonucleotide D1954.
Electron-transfer group; ETM; mismatch; genotyping;
gene expression; ss.
Synthetic.
WO200107665-A2.
01-FEB-2001.
26-JUL-2000; 2000WC-US20476.
26-JUL-1999; 99US-0145695.
17-MAR-2000; 2000US-0190259.
(CLIN-) CLINICAL MICRO SENSORS INC.
Umek RM;
WPI: 2001-159728/16.

Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface

Example 6; Page 127; 159pp; English.

The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.

Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;

Q	u	e	r	y	M	a	t	c	h	16.08; Score 278.2; DB 22; Length 936;									
B	e	s	t	L	o	c	a	l	S	i	m	i	l	a	r	i	t	y	0.9%; Pred. No. 4.3e-29;
M	a	t	c	h	7; Conservative 613; Mismatches 161; Indels 0; Gaps														

Q	y	823	aatttttcattttcgaaattgtctatacaataataataataattttaccacttcgt	882
D	b	2	www.....	61

5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843

SQ Sequence 217 BP; 66 A; 50 C; 49 G; 52 T; 0 other;

Query Match 7.5%; Score 130; DB 21; Length 217;
 Best Local Similarity 100.0%; Pred. No. 1.8e-09;
 Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aagcgatacaattccacaggaacagctatgaccatgattagcccaagcttggtacc 60
 |||||
 DB 156 ACCGGATAACAATTTTCACACAGGNAACAGCTATGACCATGATTACGCCAAGCTTGGTACC 97
 |||||

QY 61 gqctcgatccactagtaacggccgcagctgtgctggaattcggtttgtaatacagactc 120
 |||||
 DB 96 GAGCTCGGATCCACATGATACGCGCCGCGCAGTGTGCTGGATTCGGCTTGTATACGACTC 37
 |||||

QY 121 actatagggc 130
 |||||
 DB 36 ACTATAGGCG 27
 |||||

RESULT 14
 AAH93026/c
 ID AAH93026 standard; DNA; 700 BP.
 XX
 AC AAH93026;
 XX
 DT 09-OCT-2001 (first entry)
 XX
 DE Human inflammatory bowel disease related gene fragment IGR3319a.
 XX
 KW Human; inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 KW single nucleotide polymorphism; SNP; chromosome 19p13; paternity test;
 KW chromosome 5q31-33; forensic test; gene therapy; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200142511-A2.
 XX
 PD 14-JUN-2001.
 XX
 PF 11-DEC-2000; 2000WO-US33632.
 XX
 PR 10-DEC-1999; 99US-0170257.
 PR 10-APR-2000; 2000US-0196046.
 XX
 PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
 PA (ELLI-) ELLIPSIS BIOTHERAPEUTICS CORP.
 XX
 PI Daly M, Hudson TJ, Lander ES, Rioux J, Siminovitch K;
 XX WPI; 2001-367874/38.
 XX
 PT Testing for the presence of polymorphisms associated with inflammatory
 PT bowel disease, using a hybridization assay -
 XX
 PS Disclosure; Page 384; 463pp; English.
 CC
 CC The present invention describes a method for detecting the presence of
 CC polymorphisms associated with inflammatory bowel diseases such as
 CC ulcerative colitis and Crohn's disease. The methods can be used to detect
 CC the presence of genetic polymorphisms associated with inflammatory bowel
 CC disease and correlating their occurrence with disease states. They may be
 CC used in this way for phenotypic correlations, forensics, paternity
 CC testing, medicine and genetic analysis. The present sequence is a gene
 CC containing a polymorphic site described in the exemplification of the
 CC invention.
 XX
 SQ Sequence 700 BP; 370 A; 13 C; 14 G; 302 T; 1 other;

QY 904 taataaattgaatcgcatataaatttaaaattttccccaaaaaagtagactagtctca 963
 |||||
 DB 691 TATATTAAT 632
 |||||

QY 964 tctaaaaatttgattcccccaatagaaacaaattctcaaaatgaacaaacatttgaatttc 1023
 |||||
 DB 631 TAATNTAAT 572
 |||||

QY 1024 tcgatatagaagaacatttacttatttggaaatggggacacattccccaaagttttcccaaac 1083
 |||||
 DB 571 TATATTAAT 512
 |||||

QY 1084 gtaaccttgaaaggaaagttgattgagattacatccatatttttggtttttccattggaat 1143
 |||||
 DB 511 AT 452
 |||||

QY 1144 ttcattggaaaattaaaatgcacacaaaatg-----atgcatgagattaaaccaaag 1194
 |||||
 DB 451 AT 392
 |||||

QY 1195 tttatcgattgaaattcttttatttaaaaaaaccaacaaatttttaaaacttggttgcaat 1254
 |||||
 DB 391 TTTAT 332
 |||||

QY 1255 agaccaatagtttaacatcgctctattgtagataaaattgtaaatatttggtttatat 1314
 |||||
 DB 331 AT 272
 |||||

QY 1315 ttaataaattttgattttatttggatataatttggatttagatacaaaaaaattgaatttta 1374
 |||||
 DB 271 AT 212
 |||||

QY 1375 aatattttttatcatcttaataataaacatttggtaatttttttttttttttttttttt 1434
 |||||
 DB 211 TTTTAT 152
 |||||

QY 1435 ctattttttatataacatttttaataactaaatgatgtgacacacacactaat-tattttt 1493
 |||||
 DB 151 AT 92
 |||||

QY 1494 atccaaagaaaataatgctataaa 1517
 |||||
 DB 91 AT 68
 |||||

RESULT 15
 AAX59168
 ID AAX59168 standard; cDNA; 3244 BP.
 XX
 AC AAX59168;
 XX
 DT 06-SEP-1999 (first entry)
 XX
 DE Rat G protein coupled receptor flh84g5 cDNA.
 XX
 KW G protein coupled receptor; flh84g5; rat; diagnosis; screening;
 KW therapy; antiparkinsonian; nootropic; neuroprotective;
 KW neuroleptic; antidepressant; antiarrhythmic; antidiabetic;
 KW antiinflammatory; phosphatidylinositol; ss.
 XX
 OS Rattus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 778..2115
 FT /tag- a
 FT /note- "this region is specifically claimed in
 FT Claim 7b"
 XX
 PN WO928470-A1.
 XX
 PD 10-JUN-1999.
 XX

Query Match 6.3%; Score 109.4; DB 22; Length 700;
 Best Local Similarity 50.8%; Pred. No. 8.7e-07;
 Matches 317; Conservative 0; Mismatches 297; Indels 10; Gaps 2;

PF 04-DEC-1998; 98WC-US25832.
XX
PR 17-MAR-1998; 98US-0042780.
PR 04-DEC-1997; 97US-0985090.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Distefano P., Glucksmann MA, Goodearl ADJ, Xie M;
PI
XX WPI; 1999-394858/33.
DR P-PSDB; AAY06323.
DR
XX
XX New nucleic acid encoding an isolated G-protein coupled receptor
PT useful for treating nervous system related disorders
PT
XX
XX Claim 7b; Fig 2; 140pp; English.

XX This nucleotide sequence, the coding region of which is claimed,
CC codes for a novel rat G protein coupled receptor, termed
CC flh84g5 (see AAY06323). The flh84g5 cDNA was isolated from a rat
CC frontal cortex cDNA library, and has been used to isolate human
CC flh84g5 (see AAX59167). The invention provides human, rat and mouse
CC flh84g5 polynucleotides (including polynucleotides encoding the
CC transmembrane regions of flh84g5 and antisense nucleic acid
CC molecules), expression vectors, host cells, transgenic animals,
CC flh84g5 polypeptides and antibodies, and a method of modulating
CC phosphatidylinositol metabolism. The flh84g5 polypeptides can
CC (i) interact with a flh84g5 ligand, such as acetylcholine or
CC carnitine, (ii) interact with a G protein or another protein which
CC naturally binds to flh84g5, (iii) modulate the activity of an ion
CC channel (e.g. a calcium activated chloride channel or a potassium
CC or calcium channel), (iv) modulate cytosolic ion, e.g. calcium
CC concentration, (v) modulate the release of a neurotransmitter, e.g.
CC acetylcholine or carnitine from a neuron, (vi) modulate a flh84g5
CC ligand response in a responsive cell, (vii) signal ligand binding
CC via phosphatidylinositol turnover, and (viii) modulate phospholipase
CC C activity. The products can be used to treat disorders mediated
CC by abnormal flh84g5 polypeptide activity such as nervous system
CC related disorders, e.g. amnesia, apraxia, agnosia, amnesic
CC dysnomia, amnesic spatial disorientation, Kluver-Bucy syndrome,
CC Alzheimer's related memory loss and learning disability; disorders
CC affecting consciousness such as visual hallucinations, perceptual
CC disturbances or delirium associated with Lewy body dementia,
CC schitzo-effective disorders, schizophrenia with mood swings,
CC depressive illness (primary and secondary); affective disorders
CC such as REM sleep abnormalities in patients suffering from e.g.
CC depression, paradoxical sleep abnormalities, sleep-wakefulness, and
CC body temperature or respiratory depression abnormalities during
CC sleep; disorders affecting pain generation mechanisms e.g. pain
CC related to irritable bowel syndrome or chest pain; movement
CC disorders e.g. Parkinson's disease related movement disorders;
CC eating disorders e.g. insulin hypersecretion related obesity or
CC drinking disorders, e.g. diabetic polydipsia; smooth muscle related
CC disorders, e.g. irritable bowel syndrome, diverticular disease,
CC urinary incontinence, esophageal achalasia or chronic obstructive
CC airways disease; cardiac muscle disorders, e.g. pathologic
CC bradycardia or tachycardia, arrhythmia, flutter or fibrillation;
CC and gland related disorder such as xerostomia or diabetes mellitus.
CC The products can also be used for detection, diagnosis and drug
CC screening.

XX
SQ Sequence 3244 BP; 609 A; 1052 C; 903 G; 680 T; 0 other;

Query Match 6.3%; Score 109; DB 20; Length 3244;
Best Local Similarity 100.0%; Pred. No. 8.6e-07;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 agcggatacaatttcacacaggaacagctatgaccatgattacgcaagcttggtacc 60
Db 130 agcggatacaatttcacacaggaacagctatgaccatgattacgcaagcttggtacc 189
Oy 61 gagctcggatccactagtaacggccgccagtgctggaattcggcttg 109

Db 190 gagctcggatccactagtaacggccgccagtgctggaattcggcttg 238
Search completed: April 2, 2002, 09:22:39
Job time: 10390 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 2, 2002, 05:37:48 ; Search time 1546.5 Seconds
(without alignments)
12055.561 Million cell updates/sec

Title: US-09-811-093-42

Perfect score: 1735

Sequence: 1 agcgatacaaatcacc.....taattggaagcttgtagt 1735

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:**
1: em_estfun:**
2: em_esthum:**
3: em_estin:**
4: em_estom:**
5: em_estpl:**
6: em_estba:**
7: em_estro:**
8: em_estov:**
9: em_hic:**
10: gb_est1:**
11: gb_est2:**
12: gb_hic:**
13: gb_gss:**
14: em_gss_fun:**
15: em_gss_hum:**
16: em_gss_inv:**
17: em_gss_pln:**
18: em_gss_pro:**
19: em_gss_rod:**
20: em_gss_vrt:**
21: em_gss_other:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	131.4	7.6	1101	13	CNS00EVL
c 2	124.2	7.2	1101	13	CNS0021J
c 3	119.2	6.9	1201	13	CNS0167M
c 4	118.2	6.8	1101	13	CNS00EVL
c 5	117.2	6.8	625	13	CNS036A2
c 6	115.8	6.7	1101	13	CNS0021J
c 7	115.2	6.6	1225	13	CNS0161D
c 8	114	6.6	942	13	CNS00601
c 9	114	6.6	951	13	AZ676519
c 10	113	6.5	1151	11	EG309087
c 11	110	6.3	485	11	EG673765
c 12	110	6.3	638	11	BF294063

c 13	110	6.3	1101	13	CNS00B07
c 14	109	6.3	1092	13	CNS020K7
c 15	108	6.2	430	13	AQ934864
c 16	108	6.2	440	11	BG673763
c 17	108	6.2	466	10	AJ293456
c 18	107.8	6.2	1101	13	CNS00B01
c 19	107	6.2	288	11	BG319604
c 20	106.8	6.2	920	13	AZ691914
c 21	104.8	6.0	437	10	AW697868
c 22	104.4	6.0	1101	13	CNS00E07
c 23	104.2	6.0	843	13	CNS0091L
c 24	104.2	6.0	1400	5	BF627293
c 25	103.8	6.0	458	10	AL514085
c 26	103.8	6.0	1101	13	CNS0039G
c 27	103.2	5.9	578	10	BE439760
c 28	103.2	5.9	893	13	CNS013XE
c 29	103	5.9	625	10	AW682474
c 30	103	5.9	646	10	AW682722
c 31	102.8	5.9	1027	13	CNS02M50
c 32	102.4	5.9	425	10	AL514791
c 33	102.2	5.9	1143	10	AL565457
c 34	102	5.9	869	13	AF307807
c 35	101.8	5.9	1151	11	BG309087
c 36	101.8	5.9	1169	13	CNS06KHQ
c 37	101.4	5.8	335	10	AL513597
c 38	101.4	5.8	1101	13	CNS001FB
c 39	101.4	5.8	1101	13	CNS0039G
c 40	101.2	5.8	404	10	AL514087
c 41	101.2	5.8	1101	13	CNS00RAE
c 42	101	5.8	1092	13	CNS020K7
c 43	100.8	5.8	356	10	AJ278589
c 44	100.4	5.8	1013	13	CNS06RPQ
c 45	100.2	5.8	693	10	AV682300

ALIGNMENTS

RESULT: 1
CNS00EVL/c
LOCUS

DEFINITION

CNS00EVL 1101 bp DNA 04-JUN-1999
Drosophila melanogaster genome survey sequence T7 end of BAC:
BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit fly) genomic survey sequence.

ACCESSION

AL069706.1 GI:4949849

VERSION

GSS.

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Genoscope - Centre National de Sequencage :
Submitted (02-JUN-1999) Genoscope (E-mail : seqrefgenoscope.cns.fr
BP 191 91006 EVRY cedex - FRANCE)
Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs for further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammos in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

QY	1589	cattcatactgcacagctagatattactatatcaa	1623
		: : : : : : :	
Dd	1067	ATATATTAAWATAAWAAAAAAGAAWATAAWAAA	1101

RESULT	5		
CNS036A2/C			
LOCUS		625 bp	DNA
DEFINITION		Tetraodon nigroviridis genome survey sequence PUC-ORI end of clone	GSS
LOCATION		215D15 of library G from Tetraodon nigroviridis, genomic survey	
ACCESSION			
VERSION		AL229763.1	GI:7888758
KEYWORDS		GSS; genome survey sequence.	
SOURCE		Tetraodon nigroviridis.	
ORGANISM		Tetraodon nigroviridis.	
		Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;	
		Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;	
		Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;	
		Tetraodontidae; Tetraodon.	

REFERENCE
AUTHORS
1 (bases 1 to 625)
Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
TITLE
JOURNAL
REFERENCE
AUTHORS
2 (bases 1 to 625)
Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
TITLE
JOURNAL
REFERENCE
AUTHORS
3 (bases 1 to 625)
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
Genoscope.
Direct Submission
COMMENT
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetraodon>.
FEATURES
source
1..525
Location/Qualifiers

```

http://www.genoscope.cns.fr/tetraodon.
Location/Qualifiers
1. 525
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone_"215D15"
/clone_lib="G"
/note="Genoscope sequence ID : COAG215CB08SP1-end :
PUC-ori"
BASE COUNT      394 a      39 c      4 g      170 t      18 others
ORIGIN
Query Match          6.8%; Score 117.2; DB 13; Length 625;
Best Local Similarity 52.4%; Pred. No. 9e-05;
Matches 240; Conservative 5; Mismatches 213; Indels 0; Gaps 0;
QY 221 tatcttatattttagcggttacatcttcctctcctctttcttctttctttctgctgcatt 280
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 569 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 510
QY 281 ttctcccatctattctctctttactctatttttttctcttaacattgtttagattgggt 340
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 509 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 450
QY 341 aaccaaatctgattcttctctacgctctctctctctttctctctttctttcttcgctgcga 400
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 449 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 390
QY 401 ttcttcccatttgtctatcgctttttctctctttttctcttttttttttacatcgtaaaccaataa 460
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


High quality sequence stop: 1086.
 Location/Qualifiers
 1. .1151
 /organism="Hordeum vulgare"
 /cultivar="Morex"
 /db_xref="taxon:4513"
 /clone="HVSHEC000215f"
 /clone_lib="Hordeum vulgare seedling shoot EST library
 HVCNNA0003 (Etisolated and unstressed)"
 /issue_type="Seedling shoot"
 /lab_host="TJC121"
 /note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI; For
 more details on library preparation and sequence analysis
 see <http://www.genome.clemson.edu/projects/barley/> To
 order a clone see <http://www.genome.clemson.edu/orders>"
 BASE COUNT 542 a 31 c 100 g 439 t 39 others
 ORIGIN

Query Match 6.5%; Score 113; DB 11; Length 1151;
 Best Local Similarity 47.0%; Pred. No. 0.0002;
 Matches 402; Conservative 0; Mismatches 441; Indels 13; Gaps 3;

QY 793 ttgtgtatatttctattatgggtttgtgagaattttctcatcttttcgaattgttctctatc 852
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 281 TTTTGT 340
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 QY 853 aatcataataaatatttaccacttcgttatatttccgaagaagcccttaataaat 912
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 341 AATATAACCTTAATTAATAAGGGGGGGGGGACAAAACCCCCACACCCCAATAAAAA 400
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 QY 913 tgaattcgcataataatttccccaaaaagtagactatctctctcaaaaat 972
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 401 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 460
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 QY 973 ttgattcccaatagacaataattctcaaatgaacaaacatttgaaattctcgatatag 1032
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 461 AANANAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 520
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 QY 1033 aaaaactttactattttgaattgggacatatctccaaagttattccaaacgtaacttg 1092
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 521 AAAAAAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 580
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 QY 1093 aaggaaa-----agttgattgagattcacatcttctgttttccattgaaattc 1146
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 581 TATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 640
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 QY 1147 atggaataattaaaatgacacaaatgatgatgagattaaaccaaagtttatcggtatt 1206
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 641 ATTAANAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 700
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 QY 1207 gaattcttttataaaaaaaccaaaaaattttaaaaactttgtgcaatagaccacaatatag 1266
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 701 TAATTATATATANAATAATTAATTTATTTTATTAATAATTTATTAATAATTTATTA 760
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 QY 1267 ttaatccatcgctgctattgttagaataattgtaataattttgtttatattaataaatatt 1326
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 761 TAAATTATTAATAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTA 820
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 QY 1327 ttgatttatttggatatattt-gtatttagatacaaaaaataagaatttaaatatttttt 1385
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 821 TTTTGT 880
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 QY 1386 atatcttaataaacaatttggtaatttttctctattttagaccattctctctatttttat 1445
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 881 AATAAATTTAATTTATTTATTTAATAAATAATTAAGGTTAATTAATNTAATAATTAATA 940
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 QY 1446 ataacatttcaataactaaatgatgacacacactaatatttttattccaaagaaa 1505
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 941 ANA-----TAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 994
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 QY 1506 taatgctataaaaatagggtcttcttattcatcccttcattgataatttggaaaaataaaata 1565
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||

Db Qy Db Qy Db Qy Db

995 TATTANAAATTAATAAATTAANTATNATAAATTTTAAATAATTTTAAATAATTAATAAATAAAAAA 1054

1566 aaatttaattatataattcatttcattcatctaatcgacaaagctagatatattactatcataca 1625
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

1055 AAAAAATATAAAAAAATATTTTAAAAAAAATTTAAAAAAAATTCATTTAAAAAAAATTTAAAA 1114
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

1626 actttgtgtataaaa 1641
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

1115 NNAANAANAANAANAANA 1130
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
BG673765/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

EST 01-MAY-2001
905 ciliated epithelial CDNA cell library Homo sapiens CDNA 5'
mRNA sequence.
BG673765
BG673765.1 GI:113919404
human.

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 485)
Maiti,A.K., Jorissen,M. and Bouvagnet,P.
Isolation, in silico characterization and chromosomal localization
of a group of cDNAs from ciliated epithelial cells
Genome Biol. (2001) In press
Contact: Maiti AK
Laboratory de Genetic Molculaire Humaine, Faculty de Pharmacy
University Cl. Bernard,
8 Avenue Rockefeller, F69373 Lyon cedex, France
Tel: (33) 478 77 44 25
Fax: (33) 478 77 75 68
Email: amit.maiti@medecine.unige.ch
Seq primer: M13 Forward and reverse.
Location/Qualifiers
source
1. .485
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="ciliated epithelial CDNA cell library"
/tissue_type="Nasal biopsy"
/cell_type="cultured ciliated epithelial cells after in
vitro cillogenesis"
/dev_stage="Adult"
/note=Vector: PCR2.1; cloned RT-PCR fragment from
isolated total RNA"
BASE COUNT 96 a 133 c 132 g . 121 t 3 others
ORIGIN

Query Match 6.3%; Score 110; DB 11; Length 485;
Best Local Similarity 100.0%; Pred.No. 0.00066;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agccgataacaattcacagagaacagctatgccattaccgcgaagtgggtacc 60
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 227 AGCGATPACAAATTCACAGGAAACAGCTATGACCATGTACGCCAGCTTGTTGACC 168
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 61 gagctcgattccactagtaacggccgcagtgctgtcggaattcggttgt 110
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 167 GAGTCGGATCCACTAGTAGCAACGGCCGCAGTGTGCTGGAATTCGGCTTGT 118
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
BF294063/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

EST 17-NOV-2000
BF294063 638 bp mRNA
H2170-115 Human H2170 CDNA library Homo sapiens CDNA clone
H2170-115, mRNA sequence.
BF294063
BF294063.1 GI:11225127
human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 638)
AUTHORS Petersen, S., Pietas, A., Cheng, Y., Schluens, K., Pacyna-Gengelbach, M.,
Deutschmann, N., and Petersen, I.
TITLE A genome-wide survey of lung cancer associated genes
JOURNAL Unpublished (2000)
COMMENT Contact: Petersen I
Institute of Pathology
Humboldt University Hospital, Charite
Schumannstrasse 20/21, 10117 Berlin, Germany
Tel.: 00493028022611
Fax: 0049302802371
Email: iver.petersen@charite.de
Insert Length: 638 Std Error: 0.00
Seq primer: M13 Forward
POLYA=A.
FEATURES
source
1..638
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="H2170-115"
/clone_lib="Human H2170 cDNA library"
/tissue_type="non-small cell lung carcinoma (NSCLC)"
/cell_line="H2170"
/lab_host="TOP10" (ampicillin resistant)
/note="Organ: lung; Vector: pCR2.1 (Invitrogen); Site: 1;
Not1; Site: 2; SalI; The library was generated by
suppression subtractive hybridization (SSH)."
BASE COUNT 146 a 168 c 172 g 152 t
ORIGIN
Query Match 6.3%; Score 110; DB 11; Length 638;
Best Local Similarity 100.0%; Pred. No. 0.00057;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 agcggaacaatttcacaggaacagctatgaccatgattacccgaagctgtgacc 60
DB 352 acgggatacaatttcacaggaacagctatgaccatgattacccgaagctgtgacc 293
QY 61 gacgtcgatcacatgtagcagccagctgtgaggaatcggtgt 110
DB 292 gacgtcgatcacatgtagcagccagctgtgaggaatcggtgt 243
RESULT 13
CNS00E07 1101 bp DNA GSS 04-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC:
DEFINITION BACR29P01 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL069440
VERSION AL069440.1 GI:4949583
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Direct Submission
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
COMMENT - Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and

Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
ECORI digestion of Drosophila DNA and provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
source
1..1101
Location/Qualifiers
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR29P01"
/note="end : TET3"
BASE COUNT 366 a 66 c 104 g 351 t 214 others
ORIGIN
Query Match 6.3%; Score 110; DB 13; Length 1101;
Best Local Similarity 39.1%; Pred. No. 0.00044;
Matches 275; Conservative 108; Mismatches 302; Indels 18; Gaps 3;
QY 985 atagaacaatttcacaaatgaacaaacatttgaaattctcgatatagaaaacatttact 1044
DB 1101 AAAMWAAWAAAYATCMACMACACMCTWTTTTHCCCCACCCCMAMWCMCC 1042
QY 1045 tatttgaattggacataattccaaagtttattccaaacgtaactttgaaggaaggttg 1104
DB 1041 MWTTTTTTTTTTMMWAWACAMMAYTWTWTATTTTATTTTMAATATYCATAT 982
QY 1105 attgagattcacatcattttgttttttcattatgaatttcattggaataaataatgca 1164
DB 981 TTTMAATACA-----WNAAYATTTTWTATACATAATWTTTATATACAAATTTWA 931
QY 1165 cacaaatgatgatgagattaaacaaagtttctgctgtaattttttattataaa 1224
DB 930 AAATAAAWAAATTAACWAAATTAWAAWCAATTTTAAWTAATTTTATACAAWTTTA 871
QY 1225 accaacaatttttaaaactgtgttgcaatagaccataatagtagtaacatcgcgtgctca 1284
DB 870 WMTAAAAAATTTTAAWTAWAAWAAWAAWAAWTAATTTAYATWATATWTAWAAWTTATA 811
QY 1285 ttgtagataaattgttaattttgtttattttaataaattttgtttttgtttttgtatata 1344
DB 810 WATTATTTMMAATTTTATATWNAATTTTWTAAWAAWTTTWTATWAAWTAATTTTAA --- 753
QY 1345 ttgtatttagatacaaaatgaatttaaatattattattttatttatttataataaacatt 1404
DB 754 --WTTWTTTAAATTAATAAAATTTATTTTATTTTATTTTAAWAAWTTTWTTTTAAATTTW 697
QY 1405 tgttaattttttctatttttagaccattctctct---tatttttatacaacatttataaac 1461
DB 696 YTTTAATAWTTAAATTTTAAWAAWAAWAAWTAATTTAAATAAATTTTATTTAAATAATTTWAAWAT 637
QY 1462 taaatgatgtgacacacacataattatttttttccaaagaaataatgcataaataat 1521
DB 636 AAATTTAAWADWTTTATATWTTTWTTCWTTTAAATATATAAAATWAAATAATTTAAATTTT 577
QY 1522 gggctcttttaccaccttcattgataattatgaataaataaataaataaataaataaataa 1581
DB 576 TTTATATWAAWAAWAAWAAWAAWTTTATTTATTAWAAWAAWTTTATTAWAAWTTTATTTATTT 517
QY 1582 ttcatcttcatctacgtacacagctagatattactatcaacaactttgtgtataaaaa 1641
DB 516 ATTWTTTATTAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTC 457
QY 1642 gggcaagaataaagcattatcgtgagccacttttttctata 1684
DB 456 TTTCMAAWTATTWTTCCCTTTTWTMTATGTTTTTTTTTTTTTTTTH 414

Wed Apr 3 10:09:06 2002

Qy 61 gagctcggatccactagtagtaacggccgagtgctgctggaattcggctt 108
Db 334 GAGCTCGGATCCACTAGTAGACGGCCGCCAGTGTGCTGGAAATCGGCTT 287

Search completed: April 2, 2002, 07:01:09
Job time: 5001 sec

Result No.	Score	Query Match	Length	DB	ID	Description
1	109	6.3	3244	3	US-09-165-543-4	Sequence 4, Appli
2	92.8	5.3	636	4	US-08-998-416-1137	Sequence 1137, Ap
3	89.8	5.2	837	4	US-08-998-416-288	Sequence 288, App
4	89.2	5.1	7218	1	US-08-232-463-14	Sequence 14, Appl
5	87	5.0	19124	2	US-08-487-846B-13	Sequence 13, Appl
6	86.2	5.0	19124	2	US-08-487-836B-13	Sequence 13, Appl
7	84	4.8	615	4	US-08-998-416-186	Sequence 186, App
8	80.8	4.7	1431	4	US-09-316-083-2	Sequence 2, Appli
9	76.2	4.4	5852	1	US-07-867-106-2	Sequence 2, Appli
10	75	4.3	5852	1	US-07-867-106-2	Sequence 2, Appli
11	74	4.3	240	1	US-08-628-417-6	Sequence 6, Appli
12	73.6	4.2	6243	2	US-09-056-075-1	Sequence 1, Appli
13	73.4	4.2	1511	1	US-07-991-867B-8	Sequence 8, Appli
14	73.4	4.2	1511	1	US-08-107-755A-8	Sequence 8, Appli
15	73.4	4.2	1511	2	US-08-544-332-8	Sequence 8, Appli
16	73.2	4.2	658	4	US-08-998-416-595	Sequence 595, App
17	72.6	4.2	2447	2	US-08-998-416-14	Sequence 14, Appl
18	71.8	4.1	854	4	US-09-014-969-14	Sequence 534, App
19	71.4	4.1	3095	6	5231168-1	Patent No. 5231168
20	71.2	4.1	636	4	US-08-998-416-1137	Sequence 1137, Ap
21	70.8	4.1	724	4	US-08-998-416-683	Sequence 683, App
22	70.4	4.1	732	4	US-08-998-416-1036	Sequence 1036, Ap
23	69	4.0	3701	4	US-08-845-258-10	Sequence 10, Appl
24	69	4.0	3701	4	US-08-990-571-10	Sequence 10, Appl
25	69	4.0	3701	4	US-08-723-142A-10	Sequence 10, Appl
26	68.8	4.0	837	4	US-08-998-416-288	Sequence 288, App
27	68.4	3.9	6152	4	US-08-973-462-1	Sequence 1, Appli

RESULT 4
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:

Patent No. 5993827
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellem, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B.
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NH121.001Cp1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match 5.0%; Score 87; DB 2; Length 19124;
Best Local Similarity 46.9%; Pred. No. 1.7e-06;
Matches 438; Conservative 0; Mismatches 485; Indels 10; Gaps 5;

Qy 793 ttttgatatttctcattatggtttgtagaatttttccatttccgaaattggtctctac 852
Db 7157 ttttctgaatgtatataatgtagaatttttccatttccgaaattggtctctac 7098
Qy 853 aataaataaataattttaccatttctgtatatttccgaaagacccttaataaat 912
Db 7097 tctatataatataatataatataatataatataatataatataatataat 7038
Qy 913 tgaattcgatataataatttttcccaaaaagtagactatgtctatctaaaaat 972
Db 7037 ---atgttataaataatataatataatataatataatataatataatataat 6981
Qy 973 ttgattcccaatataagaacaattctcaaaatgaacaacatttgaattctcgatag 1032
Db 6980 ttttaaaaaataatataatataatataatataatataatataatataatataat 6921
Qy 1033 aaaaatttacttatttgaattgggacatatttccaaagtttattccaaagtaacttg 1092
Db 6920 aaaaaataa---taatacagaataatataatataatataatataatataat 6864
Qy 1093 aaggaaagttagtgagatcacatccatatttttttccatttgaatttccatgaa 1152
Db 6863 tatataatataatatt 6804

Qy 1153 aattaaatgcacacaaatgatgtatgagattaaac-caaagtttctcgattgaatt 1211
Db 6803 AAAATACGAAATACAAACATATAAAAGATATATATCAACGCTGTTTATATATTAATT 6744
Qy 1212 cttttataaaaaaaccaaaatttttaaaacttttgcaattggtgcaatagccaatagttat 1271
Db 6743 ATTAACATATAAT 6684
Qy 1272 ccattcggtgtattgttagataaaattgttaattttgtttatatttaataaaattttgat 1331
Db 6683 ATATTAGAGATAAACAAGAGACAAATATGTTCTCGTTCCTCTCTATATATATATCTA 6624
Qy 1332 ttattttgatataattgtatttagataa-caaaaataagatttaaaattatttttatctc 1390
Db 6623 TCATATTTAT 6564
Qy 1391 ttaataaacaattttgttaattttttctatttttagaccatttctctatttttataataac 1450
Db 6563 ATTATATTAAAGTAGTAT 6504
Qy 1451 attttaactaaatgatgtgacacacactaataatttttttccaaagaaaataatg 1510
Db 6503 TATATAATATATATATAGCATCAAAAAAATAAATGATAAATAACAGGAAAAATATA 6444
Qy 1511 ctataaaatagggctcttttttccatttccatttccatttccatttccatttccatttccatt 1570
Db 6443 TTATTAT 6384
Qy 1571 taattataatttcttctcctc--atcgacaagctagatttcttctctacatacaaac 1628
Db 6383 TTTTATGATTTTCTGATGATTTTATGCAATTTCAAAATTTTATGATATATATATATATAT 6324
Qy 1629 ttgtgtataaaaggcgaagaaataaagcatttctgtgagccatttttcttatctata 1688
Db 6323 AAGTAAAAATACACATTATATAATATATATATATATATATATATATATATATATATATAT 6264
Qy 1689 gagatagaaggttttaaaatcatgtctcatttg 1721
Db 6263 TGTTCTAT 6231

RESULT 6
US-08-487-826B-13
Sequence 13, Application US/08487826B
Patent No. 5993827
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellem, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Israelsen, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CP1
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-8268-13

Query Match 5.08; Score 86.2; DB 2; Length 19124;
Best Local Similarity 45.24; Pred. No. 2.2e-06;
Matches 519; Conservative 1; Mismatches 614; Indels 15; Gaps 5;

QY 577 aaagattttaaataacggttagtcaaatcaaacatgtatacaacaaattaaacgat 636
DB 4912 ATATATTTTATGAAAATTTTGTGTATGATATAGTTATAGTTAAAAAATAAACA 4971
QY 637 agaattgaataataacggttagttagttggtatccaaattttaaataacaaatcaaa 696
DB 4972 AGACAAATGGAACGATGAAATGTTACTGTATAGGTAATAATATATATATAA 5031
QY 697 cgatcgatcaccataacacgacgkacataccaaatctaaatgatcagcacaatat 756
DB 5032 ATGTTTATTTATCTTAAAAAGTCTCTATATAACATTAATAAATAATTTGCCCATTT 5091
QY 757 attatgcacattgtgcgaggtgttgacggaac-atttgtatattttctattatggg 815
DB 5092 ATAAATTAATTAACATACATTTACATAATGAAATTTGATTTTGTGTTTGTGATGAAAT 5151
QY 816 ttgtagaatttttccattttcgaattgttctatacaataataataattttacc 875
DB 5152 TATGGACTAATTTATATGCTGTAATGCGTCTATATAATTAATTAATTTTAAAA 5211
QY 876 acttcgttatattttcgaagagaccctttaaataaattgaattcgcatataataaatt 935
DB 5212 AATGAAAAATAACAAATAATATCTTGATTTTGTAGTCCAAATAGCTTAATAATAT 5271
QY 936 ttttcccaaaaaagtagactatgtctatctatacaaaattgttcccaatagacaatat 995
DB 5272 GGACTCATATATATATATATATCTTTTACAAACAGTAATTAATTAATTTTAAAT 5331
QY 996 tctcaaatgaacaaacatttgcgaattctcgatagaaacatttctattttgaatt 1055
DB 5332 CTTAATAGGAAATAAATAATTAATAAGATAGTGAATATTAAGTCATATATACA 5391
QY 1056 gggacatatcccaagtttatcccaacgtaactttgaaggaaagttgattgagattac 1115
DB 5392 TTTTAAAAATGTAACATAAT-TACAAATACGTAACATGATTATAGAAATAATAAGAA 5450
QY 1116 atccatatatttgttttccatattgaatttccatggaataataaataacgacacaaatgat 1175
DB 5451 TTTAATATTAAGGATAAATAATTAATAATTTTAAATTTATTTTTTATGTCATTTATGTT 5510
QY 1176 gtatgagattaaacaaagtttatcgatttgaattcttttataaaaaacaaacaaat 1235
DB 5511 ATATATATATATTAACATGATTGTTTTTGAATAATTTAAATATCATATAATAT 5570
QY 1236 tttaaaactgttttgcaatagacacataatagtttaacatcgctggtctattgtagataaa 1295
DB 5571 AATAAA-----TTAGTTAAAAATAAGTATTTTCATACAAAATACTAACTTATAAGTATA 5624
QY 1296 ttgcaaatatttctgtatatattcaataaataattttgatttattttgatattttgattag 1355
DB 5625 TCATATAATAATATATATATATATTTTATGTTTGTGTTTGTGTTGTTGTTGTTGTTGTT 5684

QY 1356 ataacaaaataagatttaaataattattttatttatttatttatttatttatttatttattt 1415
DB 5685 AAGTATATATGGGTGTTGTTTATTATATATATATATATATATATATATATATATATATAT 5742
QY 1416 tctatttttagccatttctctattttatttatttatttatttatttatttatttatttattt 1475
DB 5743 ATATTTAATTTGTTGTTATGTTCTGTTAAGATAGATATGCTTACATGCTTAAAGGTTATAG 5802
QY 1476 cacactaatattatttttttcccaagaataaataatgcctataaataatgcctatttatttattc 1535
DB 5803 TTTTTTTTTTTTTTTTGTACATATATAAAAAAATAGATAACTACAAATATGCATATT 5862
QY 1536 acctcatgataattatgaaaaataaaaaatttaattataataattcatttatttattc 1595
DB 5863 ACAAGAATAATTTGTTATATAAATATATATATATATATATATATATAAAGACATTAACATA 5922
QY 1596 tcgtacagctagatattactatatacaaacacttgcgtataaaaaaggcgaagaataataa 1655
DB 5923 TACTAATAAGGTAATTAATTTTATTATATATATATATATATATATATATATATATATAT 5977
QY 1656 gcattatcgtagcagccattttctattatctagatagatagaggttttaaatcatgctctc 1715
DB 5978 GTTTTACTCTCTGCTCTCTTTTGTATTATATAAATAAATAAATAAATAAATAAATAAATA 6037
QY 1716 taattggaa 1724
DB 6038 TATTTGGAA 6046

RESULT 7

US-08-998-416-186
Sequence 186, Application US/08998416
Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Reblschung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 36,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 186:
SEQUENCE CHARACTERISTICS:

```

RESULT      8
US-09-316-083-2/c
; Sequence 2, Application US/09316083A
; Patent No. 6280942
; GENERAL INFORMATION:
; APPLICANT: The Institute of Physical and Chemical Research
; TITLE OF INVENTION: Endonuclease
; FILE REFERENCE: PH-651
; CURRENT APPLICATION NUMBER: US/09/316,083A
; CURRENT FILING DATE: 1999-05-20
; EARLIER APPLICATION NUMBER: JP98/141861
; EARLIER FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0

```

1468	atgtgacacacactaataattatttttaccacaaagaaaataatgctataaaaataatggtgtct	1527
Qy		
1469	ttgtttatgatagttatataataacattttatgacatatcttaataataaaaattttttatt	1528
Db		
1470	tctttatcacctctcatgataattatgaaaaataaaaattttaattataataattcatt	1529
Qy		
1471	attatcccattttatataaaaaaacgcatattttcattttaaaaattttatcacatatattcata	1530
Db		
1472	tcatctaatc	1531
Qy		
1473	ttcccataatc	99
Db		

RESULT 9
US-07-867-106-2/c
Sequence 2, Application US/07867106
Patent No. 5389526
GENERAL INFORMATION:
APPLICANT: Slade, Martin B
APPLICANT: Chang, Andy C M
APPLICANT: Williams, Keith L
TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
Slime Moulds of the Genus Dictyostelium
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526/ris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA

Query Match 4.4%; Score 76.2; DB 1; Length 5852;
Best Local Similarity 44.0%; Pred. No. 8.9e-05;
Matches 469; Conservative 0; Mismatches 583; Indels 13; Gaps 3;

RESULT 10

QY	1141	aattc	aatgga	aaatt	caac	caaaaa	tgat	gat	gag	att	aa	caac	caag	ttt	atc	1200	
Db	1682	TTT	AAAA	AAAA	AAAA	AAAA	TG	TAAA	TACT	TAT	TG	TG	TAC	AT	TTT	1741	
QY	1201	gtt	atg	aatt	ctt	ttt	att	caaaa	caac	aaaa	attt	caaa	acit	gtt	ttg	caaa	1260
Db	1742	TTTT	TG	AT	G	AT	T	AAAA	AA	AA	AA	TG	CT	T	AT	CA	1801
QY	1261	ata	gtt	at	caat	ccat	gct	gct	tatt	gtt	agaa	attt	gtt	at	ttt	aa	1320
Db	1802	AAT	AAT	AT	TT	TA	TT	TAA	TAA	TC	TAC	TTTT	TTTT	TTTT	TTTT	TAA	1861
QY	1321	aat	attt	gtt	attt	ttt	gata	tata	ttt	gtt	att	taga	taca	aaaa	att	aa	1380
Db	1862	TTT	AAA	AA	TTTT	TTTT	TAT	TAG	AT	CT	CA	TAA	T	AAAA	AT	CA	1921
QY	1381	attt	tata	ctt	ca	ata	taaa	caatt	gtt	aa	ttt	ctt	ctt	attt	taga	ccatt	1440
Db	1922	AT	TTTT	TAA	TAT	GC	AAAA	CT	T	AAAA	AA	CT	TA	TG	TTT	TA	1981
QY	1441	ttt	tata	ca	catt	tt	ta	ata	ca	aat	gat	gtg	ac	ac	ac	ca	1500
Db	1982	CT	TTTT	TTTT	TTTT	TTTT	TTTT	TTTT	TTTT	TTTT	TTTT	TTTT	TTTT	TTTT	TTTT	TTTT	2036
QY	1501	gaaa	ata	tg	ct	ata	aa	at	ggg	ctt	ctt	ttt	at	cc	ctt	ct	1560
Db	2037	AAA	AA	CC	CT	CA	T	A	TAA	AT	T	AA	T	ACT	TTG	TTT	2096
QY	1561	aa	ata	aaa	ttt	aa	t	ata	ta	at	ctt	ctt	ct	aa	ct	ct	1620
Db	2097	ATT	TAA	AA	TT	TAT	TG	CT	TA	T	T	AT	CT	TAT	T	TA	2156
QY	1621	ca	aca	act	ttg	t	g	t	a	a	a	ggg	c	a	a	a	1680
Db	2157	AAT	AT	TT	T												

RESULT 11
US-08-628-417-6/c
; Sequence 6, Application US/08628417
; Patent No. 5627054
; GENERAL INFORMATION:
; APPLICANT: GILLESPIE, DAVID
; TITLE OF INVENTION: COMPETITOR PRIMER ASYMMETRIC
; TITLE OF INVENTION: POLYMERASE CHAIN REACTION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: U.S. ARMY CHEMICAL AND BIOLOGICAL
; ADDRESSEE: DEFENSE COMMAND
; STREET: OFFICE OF THE CHIEF COUNSEL (AMSCB-GC)
; CITY: ABERDEEN PROVING GROUND
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21010-5423
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/628,417
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BIFFONI, ULYSSES J
; REGISTRATION NUMBER: 39,908
; REFERENCE/DOCKET NUMBER: DAM 398-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 410-671-1158

Query Match	4.3%	Score 75;	DB 1;	Length 5852;
Best Local Similarity	44.9%;	Pre. No. 0.00014;		
Matches 372;	Conservative 0;	Mismatches 450;	Indels 7;	Gaps 2;

QY	903	ttataataaatgaattcgcataataataattttcccaaaaaagtagactatgtct	962
Ddb	1442	TTTAATAAAATAGTTTGTAGTTTAAATTTTATATCATTTTTTAAANAATGAAATGTTTGA	1501
QY	963	atctaaaaattgtattcccaatataagacaaaattctcaaaatgaacaaacatttggaaatt	1022
Ddb	1502	AAAAAAAATTTTTTTTTTTTTTTTCAACGGGACGATGAATATCATATGATTCAAAATT	1561
QY	1023	ctcgat--atagaaaaacatttacttatttggattgggacatatcccaagttttattcca	1080
Ddb	1562	AAAAGTTATTAAACAAATATCTGAANAATTATAAAAACCTAACCTAGTTATTATTCTTCC	1621
QY	1081	aacgtacctttgaaggaaaagtgttgagattacatcccatatttttgttttcatattg	1140
Ddb	1622	CCTCTTTTTTTTTTTTTTTTGTCTATGACACATTTTTTTTTTTTGTCTATGACACATTTTTT	1681

REFERENCE/DOCKET NUMBER: UF114.C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 904-375-8100

TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 1511 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:

ORGANISM: Amsacta moorei entemopoxvirus

FEATURE:

NAME/KEY: CDS

LOCATION: complement (18..218)

FEATURE:

NAME/KEY: CDS

LOCATION: complement (234..782)

FEATURE:

NAME/KEY: CDS

LOCATION: 852..1511

US-07-991-867B-8

Query Match 4.2%; Score 73.4; DB 1; Length 1511;

Best Local Similarity 46.3%; Pred. No. 0.00024;

Matches 355; Conservative 0; Mismatches 401; Indels 11; Gaps 3;

QY 824 attttttcatttcgaaattgttctatacaataataataataattttaccacttcgtt 883
 DB 421 ATATTCTGTAAAGTCACAAATTTATCCAGCACAAATACATTTTTTTTATTATTAGCC 480
 QY 884 atatttcgaagagcccttaataataattgaaattgcataataataattttttccca 943
 DB 481 ATTTTATCAAAATTTCTAAATCATTTCTTCAAAAAATGACATCATCTATGCCA 540
 QY 944 aaaaaagtagactatctatctataaaaaattgattcccaataatagacaaaattctcaaa 1003
 DB 541 ATAAATATCAAT-----TATTCAGATATTGATTTTCATTAATTAATTTTGTGTTAA 595
 QY 1004 tgaacaaacatttgaaattcctgatagaaacatttctatttgaattggggacata 1063
 DB 596 TGTATAAATATCTTTATTATATATATTTCCGTCATGATTTATATATTTTATTATATA 655
 QY 1064 ttccaaagtatttccaaacgtaactttgaaggaagaattgattgagattacatccatat 1123
 DB 656 ATCTATTATCTATATATGAGTTATATATACACATTTTGTGATTAGATAAATATATCTAT 715
 QY 1124 tttgtttttcattgaatttcattggaataataataataataataataataataataata 1183
 DB 716 TAAATTTTCCGATCAATTCGTGTGTTTCCGAGAAACATAGACCAATTTATTAATTTCTA 775
 QY 1184 ttaacacaaagtattcgttattgatttctttattataaaacacaaatttttaaac 1243
 DB 776 -TCGACATTTTTTTTATTATTGATATATTTTCAAAAAAATATATCAATCAATGAAAA 834
 QY 1244 tttgttgcattgacaaatagattgtaactcgttggtctattgttagataaaattgtaata 1303
 DB 835 AAAATAAAATATCAAAATGGATTTACTAAATTCGATATAATTTTAAATAATTTTAA 894
 QY 1304 tttgtttattataataattttgatttttattttgatatattttgatttatttagatacaaaa 1363
 DB 895 ATATATTAATTTAAAAAATAAATAAACAAGAGATATGTTTATTAATTAATATATAT 954
 QY 1364 attagatttaataattttttattctataataataacatttggtaattttttctatttt 1423
 DB 955 TAAAAAATTTAGTAAATTAGAGATTTGCATATATATATTTATGATAATAATATTTTAA 1014
 QY 1424 agaccatttcttatttttttataacatttttaataacataatgatgtgacacacactaa 1483
 DB 1015 ATAATATTTCCAGAAA-----ATATTAAAAAGTTTATATATTATTTCAATTTTAAATTTATTA 1069

QY 1484 tattattttttatccaaagaaaaataatgctataaaaaatgggtctttttttatccacttc 1543
 DB 1070 TTTAAATTTTATAACAAAATTTAAAAAATAAATATATTTAGATATATCTTTATACAAAA 1129
 QY 1544 gataaattatgaaaaataaaaaataaatttaattataataattcatttca 1590
 DB 1130 TAGCAATATAGTATATATATATATATCTACACATTTCTATAGAAATTTTAA 1176

RESULT 14

US-08-107-755A-8

Sequence 8, Application US/08107755A

Patent No. 5721352

GENERAL INFORMATION:

APPLICANT: Moyer, Richard W.

APPLICANT: Hall, Richard L.

APPLICANT: Gruidl, Michael E.

TITLE OF INVENTION: No. 5721352el Entomopoxvirus Expression System

NUMBER OF SEQUENCES: 40

CORRESPONDENCE ADDRESS:

ADDRESSEE: David R. Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: Florida

COUNTRY: U.S.A.

ZIP: 32606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/107,755A

FILING DATE: 19-AUG-1993

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/827,658

FILING DATE: 30-JAN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/657,584

FILING DATE: 19-FEB-1991

ATTORNEY/AGENT INFORMATION:

NAME: Saliwanchik, David R.

REGISTRATION NUMBER: 31,794

REFERENCE/DOCKET NUMBER: UF114.C2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (904) 375-8100

TELEFAX: (904) 372-5800

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 1511 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:

ORGANISM: Amsacta moorei entemopoxvirus

FEATURE:

NAME/KEY: CDS

LOCATION: complement (18..218)

FEATURE:

NAME/KEY: CDS

LOCATION: complement (234..782)

FEATURE:

NAME/KEY: CDS

LOCATION: 852..1511

US-08-107-755A-8

Query Match

4.2%; Score 73.4; DB 1; Length 1511;

Best Local Similarity 46.3%; Pred. No. 0.00024;

Matches 355; Conservative 0; Mismatches 401; Indels 11; Gaps 3;

QY 824 attttttcatttcgaaattgtttctatacaataataataataattttaccacttcgtt 883
DB 421 ATATTTCTGTTAAAGTCACAATTTTAAATCCAGCAACAATACTTTTATTTATATAGCC 480
QY 884 atattttcgaaagaccctcaaatcaaatgaaattcgcatataataataatttttccca 943
DB 481 ATTTATCACAATAATGTTCTAAATCAATTTCTTCAAAAATGACACTCATCTATGCCA 540
QY 944 aaaaaagtagactatgtctatacaaaaatttgattcccaataatagacaaaattccaaaa 1003
DB 541 ATAATATCATAAT-----TATCTAGCATATGATTCTTAATTAATTAATTTGTTTTAA 595
QY 1004 tgaacaaacatttgaatttcgcataatagaaacaaacttacttatttgaattgggacata 1063
DB 596 TGTATAAATATTTCTTTTAAATATATTTCCCTCATGATTTATTTATTTTATTTATAA 655
QY 1064 tcccaaaagtatttccaaacgtaacttgaaggaaaagtgtgattgagattacatccatat 1123
DB 656 ATCTATTATCTATATATGAGTTAATAATACACATTTTGTATAGATAAATAATATCTAT 715
QY 1124 ttttgttttcattgtaatttcattggaataataaaatgacacacaaaatgagtgatgaga 1183
DB 716 TAAATTTTTCGCATCAATTCGTGTTGTTTTCGCGAAGAACATAGGACCAATTAATTAATCTA 775
QY 1184 ttaacaaagtttctgttattgattcttcttatttataaaacaaacaaatttttaaac 1243
DB 776 -TCGACATTTTCTTTTATTTATTTGATATATTTTTCAAAAAATAATTAATCAATGAAAAA 834
QY 1244 tgtttgcaatagacaaatagtaattcattcgtgctgtctattgttagataaaattgtaata 1303
DB 835 AAAATAAATTTTCAAAATGGAATTTACTAAATCTGATATATTTTAAATAATTTTAA 894
QY 1304 tttgttatttataataattttgatttattttgataattttgatttagatacaaaa 1363
DB 895 AATATTATTAATTTAAAAAATAATAAATAACAGAGATAATGTTATTAATTAATTAATAT 954
QY 1364 attagatttaaatatttatttatacttataataataaaacttggatttcttctttttt 1423
DB 955 TAAAAAATTTAGTAAATTTAGAGAAATTCATATATATATATATATATATATATATTTAA 1014
QY 1424 agaccattctctatttttataataacattttaaactaaatgagtgacacacacataa 1483
DB 1015 ATAAATTTCCAGAA-----ATATTAAAGTTTATATATTTCAAATTTAAATTTATTA 1069
QY 1484 tattatttttccaaagaaaataatgctataaaaatggttcttcttttacccttcatt 1543
DB 1070 TTTAAATTTTATAACAAATTTAAAAAATAAATATACATATTTAGATATATCTTTATAACAAAA 1129
QY 1544 gataatatagaataaaataaaatttataataataatttcatttcatttca 1590
DB 1130 TAGCAATATAAGTAATATATATATATCTACCCACATTTCTATAGAAATTTTAA 1176

RESULT 15
US-08-544-332-8
; Sequence 8, Application US/08544332
; Patent No. 5935777
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Hall, Richard L.
; APPLICANT: Gruidl, Michael E.
; TITLE OF INVENTION: No. 5935777el Entomopoxvirus Expression System
; NUMBER OF SEQUENCES: 77
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gerard H. Bencen
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/544,332
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/991,867
FILING DATE: 07-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US-08/107,755
FILING DATE: 19-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 92/14818
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,585
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bencen, Gerard H.
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: UFI14.C4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Amsacta moorei entemopoxvirus
FEATURE:
NAME/KEY: CDS
LOCATION: complement (18..218)
FEATURE:
NAME/KEY: CDS
LOCATION: complement (234..782)
FEATURE:
NAME/KEY: CDS
LOCATION: 852..1511
US-08-544-332-8

Query Match 4.2%; Score 73.4; DB 2; Length 1511;
Best Local Similarity 46.3%; Pred. No. 0.00024;
Matches 355; Conservative 0; Mismatches 401; Indels 11; Gaps 3;
QY 824 attttttcatttcgaaattgtttctatacaataataataataattttaccacttcgtt 883
DB 421 ATATTTCTGTTAAAGTCACAATTTTAAATCCAGCAACAATACTTTTATTTATATAGCC 480
QY 884 atattttcgaaagaccctcaaatcaaatgaaattcgcatataataataatttttccca 943
DB 481 ATTTATCACAATAATGTTCTAAATCAATTTCTTCAAAAATGACACTCATCTATGCCA 540
QY 944 aaaaaagtagactatgtctatacaaaaatttgattcccaataatagacaaaattccaaaa 1003
DB 541 ATAATATCATAAT-----TATCTAGCATATGATTCTTAATTAATTAATTTGTTTTAA 595
QY 1004 tgaacaaacatttgaatttcgcataatagaaacaaacttacttatttgaattgggacata 1063
DB 596 TGTATAAATATTTCTTTTAAATATATTTCCGTCATGATTTATTTATTTTATTTATAA 655
QY 1064 tccaaagtttatttccaaacgtaacttgaaggaaaagtgtgattgagattacatccatat 1123
DB 656 ATCTATTATCTATATATGAGTTTAAATTAATTAATTTGATTAGATAAATAATATCTAT 715

